Overcoming the Far-Field Diffraction Limit via Absorbance Modulation

by

Hsin-Yu Sidney Tsai

M.S., Electrical Engineering and Computer Science,
Massachusetts Institute of Technology (2007)
B.S, Electrical Engineering and Physics,
National Taiwan University (2005)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011

Author .....
Department of Electrical Engineering and Computer Science
March 11, 2011

Certified by .....
Henry I. Smith
Professor
Thesis Supervisor

Accepted by .....
Leslie A. Kolodziejski
Chairman, Department Committee on Graduate Students

Overcoming the Far-Field Diffraction Limit via Absorbance Modulation

by

Hsin-Yu Sidney Tsai

Submitted to the Department of Electrical Engineering and Computer Science
on March 11, 2011, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Diffraction limits the resolution of far-field lithography and imaging to about half of the wavelength, which greatly limits the capability of optical techniques. The proposed technique with absorbance modulation aims to get around the diffraction limit by using wavelength-selective chemistry to confine light to nanoscale dimension. Absorbance modulation lithography and imaging is a near-field technique that does not require scanning of a tip in close proximity or fabrication of a physically small aperture. Near-field apertures are dynamically generated in the photochromic absorbance modulation layer (AML) with only far-field illuminations.

In this thesis, the concept of absorbance modulation is explained and in-house simulation models are discussed in detail. One-dimensional experimental demonstrations of absorbance modulation lithography achieved line exposures with widths of about one tenth of the exposure wavelength. In order to extend absorbance modulation to two-dimension, a binary diffractive-optical element that generates a focused round spot at one wavelength, aligned with the central node of a ring-shaped spot at another wavelength was designed and fabricated. Lithography and imaging results applying this diffractive optical element showed evidence of point-spread function compression in lithography and contrast enhancement in imaging.

Thesis Supervisor: Henry I. Smith
Title: Professor

Acknowledgments

I would like to thank my thesis advisor, Prof. Henry I. Smith, for his academic and personal support throughout my PhD years. Hank leads by example what attitudes scholars should take to advance science and technology. When it comes to career guidance, Hank provided encouragement, inspiration, but never hands on instructions. I enjoyed my conversations with Hank and respect him a lot.

I would also like to thank my thesis committee members, Prof. Karl K. Berggren and Prof. Peter So, for their inputs on my thesis project and support for my graduation. Having both committee members outside of the country is defenitely a unique experience. I would like to give special thanks to Prof. Rajesh Menon, who worked closely with me and contributed significantly to the work on absorbance modulation. Special thanks to Trishia Andrew, who helped us with the synthesis and characterization of new AML materials. Special thanks to Jim Daley and Mark Mondel as everyone else in the NanoStructures Lab would agree that the lab and tools would not be as productive and reliable without them. Thanks to Dr. Amil Patel for being a supportive, insightful, and cheerful officemate. I was honored and delightful to share with him many important moments in his life, including his proposal preparation and PhD defense. Thanks to Dr. Euclid Moon for technical guidance on hardware and software infrastructures in the lab. Euclid was THE MAN to go to when it comes to hands-on technical issues and getting things done. Thanks to Dr. Tim Savas for suggestions on liftoff processes. Thanks to Thomas B. O'Reilly for help with Lloyd's mirror systems. Thanks to Corey Fucetola for interesting conversations and informal coaching of presentation skills. Thanks to Jie Sun for helping me with the OAI and being a very nice officemate for my last year. Thanks to Lin-Lee Cheong for the most joyful company and enjoyable conversations in the lab and in Hank's group. Thanks to Shabnam Ghadarghadr for being my officemate and one of my role models of highly competent female engineer. Thanks to Alicia Akins for administrative support.

In addition to specific persons, I'd also like to thank a number of student groups, MIT offices, and the student support structure MIT provided. Without this many

layers of support and this many people to turn to, it would not have been possible for me to prosper through my PhD years at MIT. This institute as a whole taught me a lot, not only in academics, but also in communication, social skills, culture, politics, and even in history.

In my master's thesis, I tried to compile a full list of friends to thank, but the number of friends I'd like to mention has grown beyond two pages over my PhD career. In order to not leave out anyone or take page after pages to describe my PhD years, I'd like to let my friends, who care to read my thesis acknowledgment, know that I am really thankful to meet you. Without you, I would not have been able to cheer myself up so easily in the darkest moments in my PhD life. I'm always very proud that my friends are all phenomenal. Exceptional leaders in science, engineering, student issue, business, and various other careers inspired me to extend my boundaries, to challenge myself, and to constantly learn. Thanks to friends at Intel and IBM for introducing me to the "real world" in a very gentle and delightful way. Your inspiration and unreserved guidance has brought me to continue my career in this business. It was a pleasure to speak and work with you all.

Thanks to my friends and family in Taiwan for constant support and backup. Thanks to both of my sisters and friends who visited me all the way from the other end of earth. Your visits brought the warmest weather and blessings from Taiwan. Thanks to the prosperity of online social network and internet video conference tools for keeping me connected or even more connected than ever with my old friends. I love you all.

Thanks to Scott, who recently joined my family and share my daily life with me for the past three and a half years... We have a lifetime to share :)

Without you all, this thesis would not be possible.

Contents

List of Figures

Chapter 1

Introduction

Historically, diffraction restricts the resolution of far-field images to about half of the wavelength, which greatly limits the capability of optical imaging in nanoscale science and biology. The proposed absorbance-modulation imaging system aims to get around the diffraction limit by using wavelength-selective chemistry to confine light to nanoscale dimensions.

1.1 The Demand for High Resolution

Since the work of E. Abbe in the mid 19th century, diffraction limits the spatial resolution of far-field optics to about half the wavelength [1], which means that the size of a focused line or spot cannot be made smaller than that, and two sources with a separation closer than the diffraction limit are not differentiable. The Rayleigh criterion states that the minimum resolvable distance d_{min} of an optical system (if not limited by aberration) is

$$d_{min} = \frac{1.22\lambda}{NA} = \frac{1.22\lambda}{n \sin\theta}, \tag{1.1}$$

where NA is the numerical aperture of the lens system, n is the refractive index of the medium between the final optic and the sample, and θ is the largest angle of collection (or illumination) of the system.

For visible light, this diffraction limit is about 200 nm, but the feature sizes of interest in nanoscale science and engineering are much below that length scale. With strong incentive to pattern and image smaller features, the barrier of the diffraction limit has been pushed quite far, and these exploratory ideas has led to discovery of new science and technology at the nanoscale.

1.1.1 Choice of Illumination Source

One obvious way to lower the diffraction limit is to use photon sources with shorter wavelengths, such as the 193nm light source commonly used by the semiconductor industry, extreme-ultra-violet (EUV, wavelength $\sim$ 13nm), or x-ray (wavelength $\sim$ 0.01-10nm). Although replacing the light source seems straightforward and most direct, light sources far below the visible range are usually less available, more difficult to create proper optics for, and less compatible with living environments due to high damage to substances and water absorption. Use of an immersion medium can also reduce the effective illumination wavelength by a factor equals to the refractive index of the medium. Oil and water are the most widely used as the medium for imaging and lithography. Research on high-index-immersion lithography [2] shows that a high index, low absorbance medium gave a 1.55 numerical aperture at 193nm and can hopefully achieve an NA as high as 1.7. Nevertheless, due to physical limitations of material interactions with EM waves, resolution enhancement through engineering of the refractive index is often incremental.

Another popular way to overcome the optical diffraction limit is to use charged-particle systems, such as scanning-electron-beam systems, focused-ion-beam systems, and transmission-electron microscopy (TEM). The diffraction limit in that case is the particle wavelength of the charged particles, but that length scale is beyond the achievable spot sizes limited by aberration of the electron or ion optics. These systems have been widely used to fabricate sub-10-nm nanostructures [3] and to obtain high-resolution images down to atomic resolution with large depth of focus. Nevertheless, charged-particle imaging systems commonly face throughput limitation and image distortion problems due to the repulsive nature of charged particles. In addition,

TEM requires tedious and destructive sample preparation, while scanning-electron-beam and focused-ion-beam systems usually require sample conductivity. Focused electron/ion beams can also cause damage to devices or samples, especially for biological applications. Moreover, the vacuum environment within charged-particle systems can be hostile to samples that contain liquid, for example, living organic entities. Finally, optical systems can be more easily parallelized to achieve high throughput [4], while throughput limitations of charged-particle systems are difficult to overcome.

1.1.2 The Information Throughput

The demand for high resolution is not only about making things small and packing as much information into a given area (or volume), but also about how much information can be transfered in a given time period. Taking semiconductor manufacturing as an example, the source of information is the pattern on the mask. Via projection lithography, information on the mask is transferred into the resist on the substrate all at once. As a result, the throughput of this technique is often high and can be further increased as the line width scales down or as the wafer becomes larger. Because information or contrast of the information can be degraded during the projection process, masks are often patterned with electron-beam lithography to ensure the highest information contrast to begin with. If taking into account the time spent on fabricating the mask, we see that the information throughput is eventually still limited by the high resolution patterning step. Producing the same pattern many times is very different from producing the same amount of different patterns in terms of information. While mask-based systems are best for duplicating information, most high resolution systems, especially those that attempt to pattern or image beyond the resolution limit, are serial and maskless. As will become more obvious in following sections, the trade off between throughput and resolution is very common for information generating systems. For serial systems, resolution and information throughput are directly correlated because the smaller each information unit becomes, the more points the system has to address to cover the same area. For parallel systems, resolution is inevitably limited by diffraction and the minimum pixel size will define maximum

throughput of the system.

1.2 Lithographic Patterning beyond the Resolution Limit

Many lithographic techniques that are non-optical are capable of patterning beyond the UV-visible optical diffraction limit. In addition to patterning with high-energy photons or charged particles, contact-probe lithography [5], dip-pen lithography [6], and polymer-pen lithography (PPL) [7, 8] are also capable of fabricating arbitrary patterns in a point-wise fashion. These approaches achieve nano-scale resolution by placing sharp probes in physical contact with the substrate. Patterns are created either by physically indenting the substrate surface or by depositing molecules adhered on the tip. The resolution is, therefore, determined only by the radius of curvature at the tip. Since physically scanning a contact tip is slow, two-dimensional arrays of polymer tips were used in PPL for large-area and high-throughput patterning with nanometer registration over square-centimeter areas [7, 8, 9, 10]. This parallel process can only produce an array pattern of the same structure since each pen of the array could not be separately activated.

Many other fabrication techniques are also capable of generating features with small dimensions. However, generating small features is not the same as inputting information at the corresponding high spatial frequencies. For example, although self-assembly techniques produce feature sizes on the length scale of block-copolymer chains, the resulting patterns are regular structures, such as laminar, cylinders, or spheres, of a roughly fixed size with random or guided distributions [11]. It is the fabrication of the self-assembly template that inputs information to the final lithographic patterns [12, 13]. Another example is spatial-frequency doubling techniques using double patterning or spacer doubling. Information of the final structure is still contained mostly in the lower spatial-frequency pattern. Similarly, nano-imprint of feature sizes as small as 25 nm was demonstrated, but the technique requires a high-

resolution master template to begin with [14]. Nevertheless, capability of fabricating small features without having true high spatial-frequency information input is still useful for many applications, especially for processes that combine high-resolution pattern input techniques with pattern refinement (or duplication) techniques to enhance overall throughput. As resolution, information, and throughput are all key factors in lithography and imaging, it is important to keep in mind the nature of different high resolution fabrication or imaging techniques.

In the following sections, I will discuss lithography and imaging beyond the diffraction limit with photons in the UV-visible-IR spectrum. The discussion will be categorized by how these techniques get around the diffraction limit.

1.2.1 Contact Optical Lithography and Near-Field-Scanning-Optical Lithography

The most direct way to overcome the optical diffraction limit is to record high spatial frequency components before they are lost through propagation of light. Conventional near-field scanning optical microscopes (NSOM) can also be converted into lithography system by coupling UV light sources into the near-field aperture [4, 15, 16]. Lines as narrow as 20 nm were fabricated on a self-assembled monolayer (SAM) [4]. However, this technique often suffers from low throughput, precise gapping problems [17], and damage from the scanning aperture. Parallelization of such a system is possible with an array of NSOM apertures or tips, but fabrication and gap control remains challenging. Other examples of utilizing near-field illuminations in lithography include contact optical lithography [18] and beam pen lithography [1].

In contact lithography, a mask with patterned absorbers, such as metal, are used to generate regions of high and low light intensities. Because patterns on the mask can be made via other high resolution lithography techniques, such as scanning-electron-beam lithography or focused-ion-beam lithography, high spatial frequency patterns can be replicated via contact printing as long as the mask is placed in intimate contact with the photoresist substrate. Gratings as dense as 32-nm half-pitch were patterned

with contact lithography [19]. Nevertheless, contact lithography requires fabrication of the mask to begin with, and in order to ensure zero gap between the mask and the substrate, the mask is generally required to be conformable. Since the mask needs to be in contact with the substrate many times for repeated contact printing, mask degradation must be considered. Finally, for lithography methods that utilize optical near fields, photoresist layers with a thickness approximately equal to the finest feature size should be used to properly record near-field patterns. Patterns in such thin photoresists are not only difficult to observe but also difficult to transfer, and often require additional transfer stacks [19].

In beam-pen lithography, arrays of soft polymer-based tips are coated with a gold layer except at the point of each tip to form an array of structurally deformable NSOM tips. Because the tips are now elastic, gapping within near-field distances can be achieved by bringing the beam-pen array into contact with the substrate without damaging the substrate or array [1]. Aperture creation at the point of the tips can be achieved by placing the array in contact with an adhesive poly-methylmethacrylate (PMMA) surface or by focused ion beam (FIB) lithography for diameters as small as 50 nm. By selectively addressing backend illumination of each tip, arbitrary patterns can be exposed point-by-point with reasonable throughput and features as small as 111nm points with 400 nm illumination [1].

1.2.2 Surface Plasmon Lithography

Surface plasmons (SPs) are electromagnetic surface waves propagating along a metal-dielectric interface. Strong enhancement of the electromagnetic field in the surface via resonance amplification with photon excitation can be rather strong [20]. Plasmonic nanolithography utilizes plasmonic structures, such as bowtie apertures, to produce highly confined light enhancement through SP excitation [21, 22, 23, 24]. By fabricating these structures in a metal film directly on top of the photoresist layer, sub-diffraction limited holes of 40 nm x 50 nm were recorded in photoresist [23]. Although fabricating plasmonic structures directly on the recording substrate avoids the issue of gapping, such structures cannot be reused and can only expose isolated

dots at fixed positions. Therefore, in a more flexible scheme, plasmonic structures are placed on the tip of scanning probes and used as a direct writing head [25, 26]. Arbitrary patterns can be fabricated with this scheme, but gapping needs to be controlled carefully. One way to control gapping in this case is by placing a spacer layer between the plasmonic structures and the photoresist to provide protection at the same time [27]. Lines with a 50-nm width were patterned at a speed of 20 μm/s [27]. In reference [26], they designed a self-spacing air bearing that can fly the writing array 20 nm above a sample that is spinning at 4-12 m/s and exposed 80-nm lines in photoresists [26]. Because plasmonic structures give strong local light enhancement, the writing speed and throughput of the system is enhanced.

1.2.3 Two Photon Lithography

Resolution beyond the diffraction limit is also possible with far-field illuminations via multiphoton absorption polymerization (MAP) [28]. Instead of exposing photoresists directly with incident photons, reactions in photoresists are initiated by simultaneous absorption of two or more photons. Because the probability of multiphoton absorption is nonlinearly dependent on light intensity and single photon absorption does not expose photoresists, photoresists are selective exposured at high intensity portions, such as the focal volume and around the central peak, of the exposing beam. By exposing photoresists with absorption of multiple photons, the point-spread functions (PSFs) of the exposures are effectively reduced. In practice, because the cross-section of two-photon absorption is already low, ultra-fast pulsed lasers and sensitive chemically amplified photoresists, such as SU-8, are required for reasonable throughput [29]. Chemical nonlinearly of the photoresists also plays a major role in achieving feature sizes beyond what is possible with optical nonlinearity alone, e.g., smaller than one tenth of the wavelength of the exposure light [30]. Fabrication of a woodpile photonic-crystal structure with resolution as fine as 65 nm ($\lambda/8$) was demonstrated using light of a shorter wavelength (520 nm rather than the most common $\sim$ 800 nm light) [31]. Finally, MAP is a true 3D fabrication technique since the focal point of the laser beam can easily be scanned along three axes through the photoresist volume

to fabricate 3D structures point-by-point [32].

Experimentally, Kawata et al. demonstrated fabrication of a 3-D bull sculpture and creation of 120 nm features with 800 nm light ($\lambda/6.7$) in 2001 [32]. More recently, 80 nm features ($\lambda/10$) was created with the same wavelength [33]. Two-photon lithography can also be combined with near-field scanning optical lithography and plasmonic lithography to further enhance spatial resolution and boost up two-photon reaction rates. Lithographic features as small as 70 nm were produced with $\sim$ 800 nm illumination on an atomic-force microscope (AFM) tip [34]. Isolated lines as narrow as 24 nm were achieved by combining nanoscale bowtie apertures on a scanning probe with two-photon lithography [35].

1.2.4 Nodal Lithography

The idea of nodal lithography was first proposed in 1996[36] and demonstrated with neutral-metastable-atom lithography in 1998[37]. Recently, several lithography techniques, including absorbance modulation, utilize nodal schemes to achieve resolution beyond the optical diffraction limit by limiting exposures to the immediate vicinity of the optical nodes.

The key idea is that if one can activate reaction exclusively at optical nodes, instead of at peaks of the light pattern, the size of the activated region can decrease well beyond the diffraction limit. One way to achieve activation only at optical nodes is by deactivating most activated region except where the deactivating beam has zero intensity, *i.e.* the node. Because the deactivating illumination does not cause permanent effects to the substances of interest, the only region that undergoes permanent change, such as polymerization, is the activated region. By scanning the illumination patterns in space, small features can be placed close together and high-resolution imaging or patterning can be achieved in a serial manner. With nodal lithography, neither the feature sizes of each exposure nor the distance between patterns will be bound by the Rayleigh criterion.

Many of the recent lithography techniques derive their inspiration from an imaging technique, called stimulated emission depletion (STED) microscopy [38], which will be

described in more detail in the next section. STED utilizes ground state excitation and stimulated emission of specific fluorescent molecules as its activation and deactivation mechanism, which are ideal for imaging, but not for lithography. In order to apply the concept to lithography, activation and deactivation mechanisms that apply for lithography must be identified.

Neutral-Metastable-Atom Lithography

Metastable argon atoms are capable of damaging the self-assembled monolayers[39] or prompting the deposition of carbonaceous film from vapor-phase hydrocarbon precursor upon collision with the surface[37], while ground-state argon atoms are not. An IR photon can bring the metastable argon atom to its excited energy state, which then radiatively decay to its ground state. Using a one-dimensional standing wave created with an IR laser, optical quenching of the metastable argon atoms occurs except at the nodes of the standing wave. As a result, only atoms that pass through the nodes of the IR beam remain in their metastable states and interact to form carbonaceous structures on the substrate. This nodal aperture can be made much smaller than the wavelength of the IR light. What limits the feature size of the lines, in this case, is the Heisenberg uncertainty limit of the atoms. Line patterns with 65nm FWHM, which is about $\lambda/15$, was achieved with this technique.

Absorbance Modulation Lithography

Absorbance modulation utilizes photo-switching of a photochromic film between its opaque and transparent states [40]. For lithography, a photoresist layer is placed underneath the photochromic absorbance-modulation layer (AML). When the AML is transparent (or less absorbing,) the system is "active" because the AML permits penetration of the patterning light and, therefore, the exposure of the underlying photoresist. When the AML is opaque (or highly absorbing,) the system is "deactivated" because the AML prevents light from exposing the underlying photoresist. Activation is achieved via illumination of a shorter-wavelength light which also exposes the photoresist, while deactivation is achieved via illumination of a longer-wavelength

light that does not affect photoresists. Unlike other STED-like schemes, where the activated (deactivated) material is the substance of interest, in absorbance modulation, activation (deactivation) occurs in the AML while exposures take place in the separate photoresist layer. This separation facilitates the search of AML candidates since the AML and photoresists can be optimized separately. The photoresists is exposed with near-field illuminations leaking from the AML overlayer, so the technique becomes solely 2D. This is by no means an intrinsic limitation to absorbance modulation since the AML can also be altered to act directly as the resist. In Chapters 2-5 of this thesis, numerical analysis and experimental demonstration of absorbance modulation with two material systems will be discussed in detail.

Resolution Augmentation through Photo-induced Deactivation Lithography

Another STED-like technique for lithography is called resolution augmentation through photo-induced deactivation (RAPID) lithography. RAPID lithography utilizes one laser beam to initiate multiphoton absorption polymerization (MAP) in a negative-tone photoresist and a second laser beam at the same wavelength to deactivate the photoinitiator [41]. Because typical photo-initiators produce polymerization only when excited to the triplet state, deactivation can be achieved by de-exciting enough photo-initiators, such that the triplet-state concentration is below threshold before photo-polymerization occurs. However, dominant stimulated-emission-depletion interaction requires strong oscillator strength between the ground and first excited state [42], which is not true for most radical photoinitiators. The work in references [41, 43] identified dyes that can act as photoinitiators for MAP upon femtosecond-pulse illumination and be deactivated with considerably longer pulses or continuous wave (CW) illuminations of the same wavelength. Two-photon initiation dominates over one-photon deactivation when short excitation pulses are illuminated. In contrast, longer deactivation pulses or CW illuminations can have considerably greater overall energy than that of the excitation beam without causing polymerization. By properly shaping the initiating and deactivating beam, a 40-nm minimum feature size was

achieved with 800-nm illuminations.

Combined Use of Photo-initiators and Photo-inhibitors

A third STED-like lithography technique utilizes photoinitiators and generation of localized inhibitors, such as radical traps, for activation and deactivation [44]. This method is similar to RAPID lithography in the sense that deactivation is also achieved via deactivation of photoinitiators. However, initiation and the inhibition of radical photo-polymerization are uniquely correlated to single-photon absorption at two distinct wavelengths, i.e. initiating species are generated at one wavelength while inhibiting species are generated at a second, independent wavelength. In this case, inhibiting species are noninitiating radicals that couple with the growing polymer chain and terminate polymerization [44]. Because single-photon absorption has much larger interaction cross-section, patterning throughput is higher compared to two-photon lithography and CW diode lasers can by used. By appropriately shaping and overlapping the activation and deactivation beams, 64-nm dots on a 2D plane were fabricated [44].

1.3 Imaging beyond the Resolution Limit

The development of optical microscopy enabled tremendous discoveries in biology, physics, and semiconductor electronics. However, as diffraction limits the spatial resolution of far-field optical imaging to about half the wavelength [45], *i.e.*, $\sim$ 200 nm for $\lambda \sim 400$ nm, it limits the applicability of optical imaging techniques to nanoscale features below 100nm, which are of great interests for both nano-science and biological studies. The use of visible or infrared light sources is especially desired for biological imaging, as most substances react strongly and specifically with photons at that wavelength range and no or minimal damage is done to organic molecules. Traditionally, nanoscale imaging or nanoscopy was carried out via electron and X-ray microscopy. These high-energy particles impart considerable damage to fragile organic samples. Moreover, the vacuum environments within such systems are hos-

tile to many materials. Photons also play a key role in imaging inorganic samples [46, 47]. Specifically, they provide contrast mechanisms that are unavailable with other particles.

1.3.1 Recent Advancements in STED

Optical nanoscale imaging of biological samples is a fairly new field [48]. The major advance in this field was achieved by Hell and co-workers with the implementation of stimulated-emission-depletion (STED) microscopy [38, 49] and related techniques, such as ground state depletion (GSD) [50, 51]. In STED, a focused pulse of light is used to excite fluorescence, while a second pulse, shaped like a ring in 2D (or a hollow sphere in 3D), is used to de-excite the fluorescence everywhere except at an arbitrarily small region near the optical axis (or at the central intensity null). Using this and related techniques, it was possible to image biological samples with resolution below 20 nm [52]. Nevertheless, such techniques require specific fluorophores and relatively high light intensities in the ring-shaped pulse. These requirements arise from the metastable nature of the fluorescent (excited) state. In particular, high intensities and also high power, often achieved using short pulses, are required to stimulate non-radiative transitions to the ground state before emission occurs. Parallelization of such approaches is therefore difficult. Recent advances in STED allow for lower intensities [53], but specialized switchable fluorophores are required.

1.3.2 Stochastic Imaging of Single Molecules

A different set of techniques for sub-diffraction-limited imaging takes advantage of growing capabilities in single molecule imaging and precise position localization of molecules. Given that enough photons are detected, localization precision can be turned into resolution by temporally separating fluorescence signals from closely spaced molecules, so that they can be localized separately. Initial implementations of this concept included photoactivated localization microscopy (PALM) [54, 55, 56, 57], and stochastic optical reconstruction microscopy (STORM) [58, 59, 60]. These tech-

niques rely on stochastic switching of special fluorescent molecules between their fluorescent and non-fluorescent states. By photo-activating sparse distributions of the fluorescent molecules at one wavelength and then exciting fluorescence at a different wavelength, multiple fluorescent photons are collected from each active fluorophores before they return to the non-fluorescent state. Assume photo-activation is sparse enough, such that photon clusters from two distinct fluorophores are separable on the imaging plane, the centroid position of the collected photons represents a more accurate estimate of the true position of the single fluorophore. The accuracy of localizing a single emitter improves proportionally to the square root of the number of photons collected and can be as accurate as 1.5 nm [61]. Images are then generated by tens of thousands of imaging and activation cycles, followed by post-processing of the collected data. Because a large number of imaging cycles are required to probe all the fluorophores within the field, imaging speed and time resolution is usually limited for this type of approache. In addition, these methods rely on fluorophores being photoactivatable, which is not the case for most natural fluorophores. Nevertheless, stochastic imagining systems can be easily adopted because they require little modification to a conventional optical setup, mostly only changing the camera and fluorophore. PAML/STORM can also accomplish 3D-imaging by adding astigmatism into the imaging system [58], by defocusing light [62], by using a double-helical point-spread function [63], or by using interferometry [64]. A resolution of 10-20 nm in all three dimensions was achieved [63].

Finally, for many biological systems with smaller molecules of interest, it is always a concern that fluorescent tagging may affect the original biological system. A combination of stochastic imaging with GSD, called ground-state depletion with individual molecule return, addresses this issue by directly applying the concept to auto-fluorescence [65]. As many substances in biological tissue display natural autofluorescence, the technique opens up possibilities for label-free optical imaging beyond the resolution limit.

1.3.3 Structured Illumination

Structured-illumination microscopy (SIM) is a microscopy technique that applies frequency-space convolution principles to achieve sub-diffraction-limited imaging. By convolution theorem, point-wise real-space product of the sample and the illumination intensity corresponds to convolution of their spatial frequency information in frequency space [66]. The statement that resolution of imaging systems is limited by diffraction is equivalent to stating that imaging systems have a low spatial-frequency detection passband which is limited by the wave vector of the illuminating light. By illuminating samples with structured-illumination, for example, a sinusoidal grating illumination, some high-frequency components can be shifted into the detection passband, and information that is originally outside the passband can be observed [67, 68]. By observing the sample with a series of periodic illumination patterns, specific regions in frequency space are translated into the detection passband. Three-dimensional linear SIM was shown to produce reconstructions of otherwise unresolvable cellular structures [69].

In SIM, the highest spatial frequency present in the illumination determines how wide the passband of the imaging system can be enlarged. Unfortunately, the highest spatial frequency that illumination source can contain is itself limited by diffraction. This limits the resolution of SIM with linear photo-responses to about two times the diffraction-limited resolution. A nonlinear photo-response is required to overcome this twofold limitation. Substances with nonlinear photo-responses generate higher harmonics of the illumination frequencies, which can contain much higher spatial-frequency components. Because effective illuminations generated through nonlinear photo-responses usually also contain lower spatial-frequency components, properties of the nonlinear photo-response and signal-to-noise ratio limit the experimentally achievable resolution. To achieve theoretically unlimited resolution with nonlinear SIM, the nonlinearity must be non-polynomial [70]. A form of nonlinear SIM that uses saturation of the excited state as the nonlinear response has demonstrated 50-nm resolution [71], and other nonlinearities could be used in the same manner.

1.3.4 Multi-photon Microscopy

Three-dimensional two-photon or three-photon fluorescence imaging is attractive in biological imaging for its high spatial and temporal resolution [72]. One important reason is its capability to image several hundreds of microns deep into tissues [73]. Although MW/cm^2 to GW/cm^2 power is still required due to relatively low absorption cross-section (compared to single-photon absorption), the power requirement is usually lower for imaging applications and molecules with larger cross-sections have been widely studied [74, 75].

1.3.5 Near-Field Scanning Microscopy

As previously mentioned for lithography, utilizing near-field illuminations with optical contact probes, such as in near-filed scanning optical microscopy (NSOM) [76], can also achieve spatial resolution beyond the far-field diffraction limit. In conventional NSOM, a near-field probe (either an optical fiber with a subwavelength aperture or a cantilever with a sharp metal tip at one end) is brought within a few nanometers of the sample. Unlike in lithography, where near-field illuminations are always used to induce exposure, subdiffraction-limited resolution in imaging can be achieved by either illuminating samples with the near-field probe or by collecting optical near-field signals through the near-field probe. Similar to near-field lithography, NSOM is usually slow due to the scanning of a single tip across the sample and requires exquisite control of the gap between the probe and the sample because signal depends exponentially on the gap. Gap control becomes problematic in some applications, and serious damage can arise when the probe accidentally runs into the sample. Parallelization of NSOM is, therefore, challenging.

1.3.6 Plasmon Enhanced Imaging

As described in section 1.2.2, surface plasmons (SP) can be used alone or be combined with other techniques to enhance resolution for imaging. How SP enhances resolution can be understood in many ways. Excitation of highly-localized light with strong

intensity enhancement is one way to think about it. Using localized SP to enhance resolution and throughput of near-field scanning imaging techniques or fluorescent scanning imaging techniques is very similar to what we described for SP lithography. With a near-field scanning probe that induce SP at the apex of the tip, about 20-nm resolution was achieved with two-photon imaging [77].

The burden of scanning an SP enhancement tip across imaging samples can be avoided by placing plasmonic materials or structures on the substrate before placing the specimen of interest on top. For example, plasmon-enhanced total internal refraction fluorescence (PE-TIRF) imaging using SP from gold thin films has been of significant interest for enhancing signal strength of traditional TIRF microscopy (TIRFM) [78, 79, 80]. In some other examples, such as plasmonics-based spatially activated light microscopy [81], SP from subwavelength structures are utilized to enhance spatial resolution. By sequentially exciting SPs on each side of a grating ridges features spaced by the size of the ridge can be resolved.

Another way to understand SP effects is by examining the dispersion curve and explaining resolution enhancement as a result of slowing the speed of light near plasmonic resonance [82]. When light is slowed down, the effective wavelength of light is reduced and super-resolution can be achieved. When approaching resonance, the so called "superlens" can form perfect images without any blurring [83]. Superlens can also be explained through negative refraction [84, 85, 86], but imaging with superlenses requires both the object and the image to be in the near field. The superlens concept was first demonstrated experimentally with a 35-nm-thick silver film as the superlens and photoresist exposures to record the image plane, resulting in about 65-nm resolution [86]. Superlenses with magnifying function were also demonstrated [87, 88, 89], but fundamental properties of SP, such as near-field requirement of the object and absorption of plasmons in metal, limits the application and ultimate resolution of this technique.

1.3.7 Absorbance Modulation Imaging

The proposed absorbance-modulation imaging (AMI) technique in this thesis is a far-field technique that does not require scanning of a tip in close proximity or fabrication of a physically small aperture. Near-field apertures are dynamically generated in the photochromic absorbance modulation layer (AML) with far-field illuminations. Because AMI can be readily made massively parallelized for high throughput [90] and does not require gap control between the near-field probe and the sample, it eliminates major constraints in conventional NSOM. Because AMI separates the probe formation and signal generation processes, AMI does not rely on specific fluorescent markers and high power depletion beam, as in STED. More specifically, intensities of the continuous wave lasers used in our experiments was only 1-10 mW/cm^2, very low compared to the MW/cm^2 to GW/cm^2 intensities used in STED [91]. As a result, AMI is capable of imaging a wider range of samples, particularly those that are easily damaged by high-intensity illumination and those that cannot be easily made fluorescent.

It is worth noting that one key disadvantage of AMI compared to other far-field high-resolution techniques, such as confocal, two-photon, STED, or PALM/STORM [49], is that only a thin layer close to the surface can be imaged. This limitation can be disadvantageous for some biological applications, especially those that require imaging through a whole cell. Nevertheless, many interesting biochemical reactions happen on cell membranes, which can be observed with AMI.

1.3.8 Organization of Thesis

In Chapter 2, the concept of Absorbance Modulation and the simulation model we built will be described. Simulation results based on experimentally characterized material parameters is discussed as well as the numerical properties of the model. In Chapter 3, previous experimental work on absorbance modulation lithography for patterning grating (line) structures will be reviewed and the latest improvements made based on past experiments discussed. In Chapter 4, the design and fabrication of

a diffractive-optical element will be discussed in preparation to extending absorbance modulation from patterning grating lines to patterning round-shaped spots. This diffractive-optical element generates a focused round spot at one wavelength, aligned with the central node of a ring-shaped spot at another wavelength. In Chapter 5, we show lithography and imaging experiments that apply the diffractive-optical element designed in Chapter 4.

Chapter 2

Theory and Simulation for Absorbance Modulation

2.1 How Absorbance Modulation Works

The absorbance-modulation layer (AML), which is a photochromic film, is rendered opaque by illumination at λ_2 and transparent by illumination at λ_1, as illustrated in Fig. 2-1(a). By absorbing λ_1 and λ_2 photons, the AML reversibly converts between its two states: A (opaque) and B (transparent). Relative concentrations of A and B in the AML depend on local λ_1, λ_2 light intensities and determine the overall absorbance property of the AML. With proper design of the λ_1, λ_2 illuminations, for example, illumination of a ring-shaped spot at λ_2 coincident with a focused spot at λ_1, a laterally compressed transmitted intensity profile, *i.e.*, a compressed PSF at λ_1, can be formed through the AML. In analogy with other work utilizing similar illumination patterns, the λ_1 illumination serves as the activation beam, while λ_2 illumination provides deactivation and confinement of the λ_1 illumination. The size of the aperture thus formed is limited only by the AML photokinetic parameters and the intensity ratio of the two illuminating wavelengths [92], not by diffraction. One example of the AML consists of a polymer containing azobenzene side chains that exhibit markedly different absorbance at $\lambda_1 = 405$ nm in its two isomer forms as shown in Fig. 2-1(b).

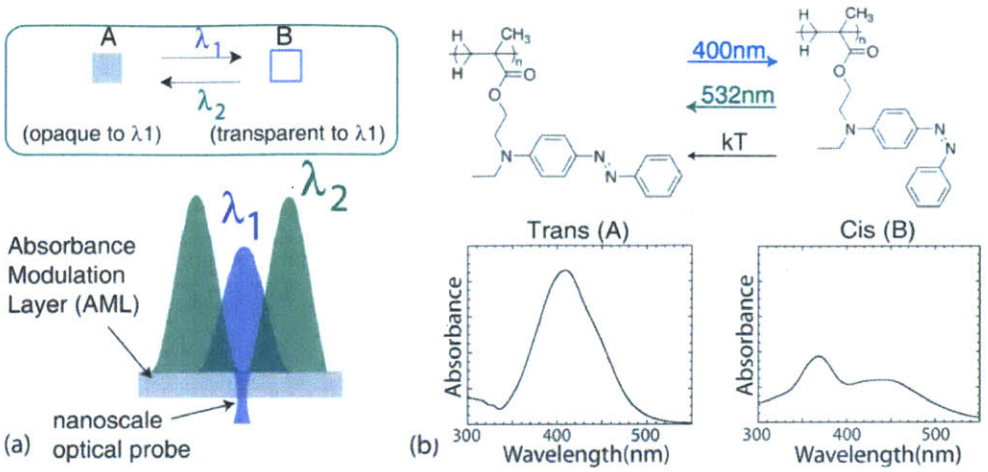

Figure 2-1: (a) Concept and illumination configuration for absorbance modulation. Through dynamic competition of the reversible transitions in the AML, a subwavelength aperture is created for λ_1. (b) AML is composed of a polymer containing the photochromic azobenzene side chain, shown at the top. Upon exposure to 405 nm light, the trans isomer undergoes photoisomerization forming the cis isomer. The reverse reaction is favored upon exposure to 532 nm light or thermal excitation. The absorbance of the two isomers at $\lambda_1 = 405$ nm is markedly different.

2.2 Theory for Absorbance Modulation

2.2.1 Model and Simulation

One way to describe light propagation through the AML is by viewing light as photons propagating through the absorptive medium. The equation that governs photon propagation is the photon transport equation [93]:

$$\frac{1}{v}\frac{\partial L(r,\hat{\Omega},t)}{\partial t} + \nabla \cdot L(r,\hat{\Omega},t)\hat{\Omega} + \mu_t L(r,\hat{\Omega},t) = \mu_s \int L(r,\hat{\Omega},t) f(\hat{\Omega},\hat{\Omega}') d\hat{\Omega}' + S(r,\hat{\Omega},t), \tag{2.1}$$

where L is the radiance with units of $\mathrm{W/m^2/sr}$ at position r, traveling in direction $\hat{\Omega}$ at time t, f is the normalized phase function that represents the probability of scattering into a direction $\hat{\Omega}'$ from direction $\hat{\Omega}$, v is the speed of light in the medium, $\mu_t = \mu_s + \mu_a$ is the transport coefficient, where μ_s is the scattering coefficient and μ_a is the absorption coefficient. S is the spatial and angular distribution of the source with units of $\mathrm{W/m^3/sr}$.

If we ignore scattering and photon sources in the AML, the terms on the right of the equation vanishes and the transport coefficient μ_t becomes equal to the absorption coefficient.The photon transport equation reduces to:

$$\frac{1}{v}\frac{\partial L(r,\hat{\Omega},t)}{\partial t} + \nabla \cdot L(r,\hat{\Omega},t)\hat{\Omega} + \mu_a L(r,\hat{\Omega},t) = 0 \tag{2.2}$$

The equation further reduces to the following equations when photons are assumed to travel only in the z direction, *i.e.*, $L(r,\hat{\Omega},t) = I(r,t)\delta(\hat{\Omega} - \hat{z})$.

$$\frac{\partial I_1(\rho,z,t)}{\partial z} + \frac{1}{v}\frac{\partial I_1(\rho,z,t)}{\partial t} = -\ln(10)(\epsilon_{1A}[A](\rho,z,t) + \epsilon_{1B}[B](\rho,z,t))I_1(\rho,z,t) \tag{2.3}$$

$$\frac{\partial I_2(\rho,z,t)}{\partial z} + \frac{1}{v}\frac{\partial I_2(\rho,z,t)}{\partial t} = -\ln(10)(\epsilon_{2A}[A](\rho,z,t) + \epsilon_{2B}[B](\rho,z,t))I_2(\rho,z,t) \tag{2.4}$$

where I_1 and I_2 are the intensities for illumination at wavelength λ_1 and λ_2, as a function of radial distance ρ from the optical axis, depth z into the AML, and time t (unit=$J/s/m^2$); [A] and [B] are the concentration of the two AML states (unit=$mole/m^3$); and ϵ_{1A}, ϵ_{1B} are the extinction coefficient (absorption coefficient) at wavelength λ_1 for AML state A, B (unit=$m^2/mole$); and similarly for ϵ_{2A} and ϵ_{2B}.

The rate equation describing photo-switching between two states of the AML gives:

$$-\frac{\partial[A]}{\partial t} = [A]I_1\epsilon_{1A}\frac{\phi_{1AB}}{E_{M1}} + [A]I_2\epsilon_{2A}\frac{\phi_{2AB}}{E_{M2}} - [B]I_2\epsilon_{2B}\frac{\phi_{2BA}}{E_{M2}} - [B]I_1\epsilon_{1B}\frac{\phi_{1BA}}{E_{M1}} - [B]k_{BA} \tag{2.5}$$

where ϕ_{1AB} is the quantum efficiency of conversion of the AML from state A to state B with illumination at λ_1 (unitless); similarly for other quantum efficiencies. E_{M1} is the energy of 1 mole of photon at λ_1 (unit=$J/mole$) and k_{BA} is the thermal conversion rate of the AML from state B to state A (unit=$1/s$). State A is assumed to be thermally stable.

Finally, the conservation of photochromic molecules in the AML, ignoring photo-

Table 2.1: Photo-kinetic parameters for the azobenzene AML used thorughout this chapter.

λ_1	405e-9 (m)	λ_2	532e-9 (m)
ϵ_{1A}	30000 $(m^2/mole)$	ϵ_{1B}	12500 $(m^2/mole)$
ϵ_{2A}	1250 $(m^2/mole)$	ϵ_{2B}	3000 $(m^2/mole)$
ϕ_{1AB}	0.0037	ϕ_{1BA}	0.0021
ϕ_{2AB}	0.0021	ϕ_{2BA}	0.0072
A_0	400.024 $(mole/m^3)$	k_{BA}	0.002 $(1/s)$

bleaching effects, gives

$$[A]_0 = [A] + [B] \tag{2.6}$$

where $[A]_0$ is the initial concentration of photochromic molecures in the AML. All photo-kinetic parameters in this work were experimentally characterized. The AML that was extensively studied was the one with azobenzene side chains, as shown in Fig. 2-1(b). Table 2.1 shows the photo-kinetic parameters for the azobenzene AML. The AML film thickness in z was chosen to be 200 nm to match experimental conditions.

2.2.2 Fast Transport Approximation

The most important assumption made in this problem is that the two time scales in the problem – photon transport time through the AML and AML absorbance response time – are drastically different. The assumption can be verified by plugging in the photo-kinetic parameters above, assuming reasonable I_1 and I_2 values. When the photon transport term is set to zero, *i.e* assuming photon transport has reached steady state, the shortest response time with the azobenzene AML is 1 ms. Since photon transport time, estimated with the thickness of the AML and the speed of light in vacuum, is only 6.67×10^{-16} s, this assumption should be valid. Therefore, the photon transport time derivative term in Eq.2.3 and Eq.2.4 can be set to zero. Substituting Eq.2.6 into Eq.2.3 to Eq.2.5, we get the following system of equations

$$\frac{\partial I_1(\rho,z,t)}{\partial z} = -\ln(10)[(\epsilon_{1B}-\epsilon_{1A})[A](\rho,z,t)-\epsilon_{1B}[A]_0(\rho,z,t)]I_1(\rho,z,t), \quad (2.7)$$

$$\frac{\partial I_2(\rho,z,t)}{\partial z} = -\ln(10)[(\epsilon_{2B}-\epsilon_{2A})[A](\rho,z,t)-\epsilon_{2B}[A]_0(\rho,z,t)]I_2(\rho,z,t), \quad (2.8)$$

$$-\frac{\partial[A]}{\partial t} = [A][I_1(\alpha_{1A}+\alpha_{1B})+I_2(\alpha_{2A}+\alpha_{2B})+k_{BA}]-[A]_0(I_1\alpha_{1B}+I_2\alpha_{2B}), \quad (2.9)$$

where $\alpha_{1A}=\epsilon_{1A}\frac{\phi_{1AB}}{E_{M1}}$, $\alpha_{1B}=\epsilon_{1B}\frac{\phi_{1BA}}{E_{M1}}$, $\alpha_{2A}=\epsilon_{2A}\frac{\phi_{2AB}}{E_{M2}}$, and $\alpha_{2B}=\epsilon_{2B}\frac{\phi_{1BA}}{E_{M2}}$.

2.2.3 Steady State Approximation

In addition to the fast transport approximation, which is essential for the following numerical analysis, some additional approximations can be made to further simplify the system and obtain useful estimates of the outcome. One very useful simplifying assumption is the steady state approximations.

In steady state, Eq. 2.9 can be further simplified to directly associate $[A]$ at each point with its local I_1, I_2 intensities as follow:

$$A_{steadystate} = \frac{\alpha_{1B}I_1+\alpha_{2B}I_2+k_{BA}}{(\alpha_{1A}+\alpha_{1B})I_1+(\alpha_{2A}+\alpha_{2B})I_2+k_{BA}}A_0 \quad (2.10)$$

In that case, the steady state intensities within the AML can be solved by substituting $[A]$ in Eq. 2.7 and Eq. 2.8 with Eq. 2.10 and solving the system of differential equations for I_1 and I_2. The Matlab code for solving steady-state distibutions are documented in Appendix A. The ODE solver used was the 'ode15s' function in Matlab. While the 'ode15s' function provides only low to medium accuracy, it runs much faster than the most commonly applied 'ode45' solver.

Steady-state approximation is very powerful and is usually valid for our experimental conditions, especially for experiments conducted with the azobenzene AML (Fig. 2-1) because its photo-transitions are fast compared to exposure time. The different AMLs and their photokinetic properties will be discussed in more detail in Chapter 3. Since the steady-state simulator was most widely applied throughout this

work to compare theory with experimental results, the Matlab code was developed to adapt four kinds of illuminations:

- Analytical expression for focused spots with a planewave phase front at λ_1 and a spiral phase front at λ_2,
- focal plane λ_1 and λ_2 profiles for a two-wavelength diffractive-optical element, called a dichromat (explained in more detail in Chapter 4),
- standing wave illumination at λ_2 with flood illumination at λ_1, and
- standing wave illuminations at both λ_1 and λ_2.

In the following sections, we will focus discussion on simulation results with the first two illumination types and their transient responses. Since the input illuminations only set the boundary conditions $I_1(\rho; z = 0)$ and $I_2(\rho; z = 0)$ for the steady-state problem and set the initial conditions $I_1(\rho; z = 0; t)$ and $I_2(\rho; z = 0; t)$ for transient analyses, the same simulators can be easily extended for additional illumination types.

2.3 Full-Width-Half-Maximum (FWHM) and Efficiency scaling at steady state

The first kind of I_1, I_2 profiles is generated with dichromats. Dichromats generate round-shaped spots at λ_1 and a doughnut-shaped spots at λ_2 at the focal plane of the element. The design and experimental validations of dichromats will be described in detail in Chapter 4. The focal plane intensities used as the boundary conditions were computed with scalar diffraction theory. Typical solutions for two different λ_2 to λ_1 ratios are shown below in Fig. 2-2 for an NA=0.7 dichromat input. The left figures are with $I_2/I_1 = 200$ and right figures with $I_2/I_1 = 20000$. The cross-sectional profiles and output intensity profiles show that higher I_2/I_1 ratios result in greater compression of the input λ_1 profile and, therefore, a smaller output FWHM.

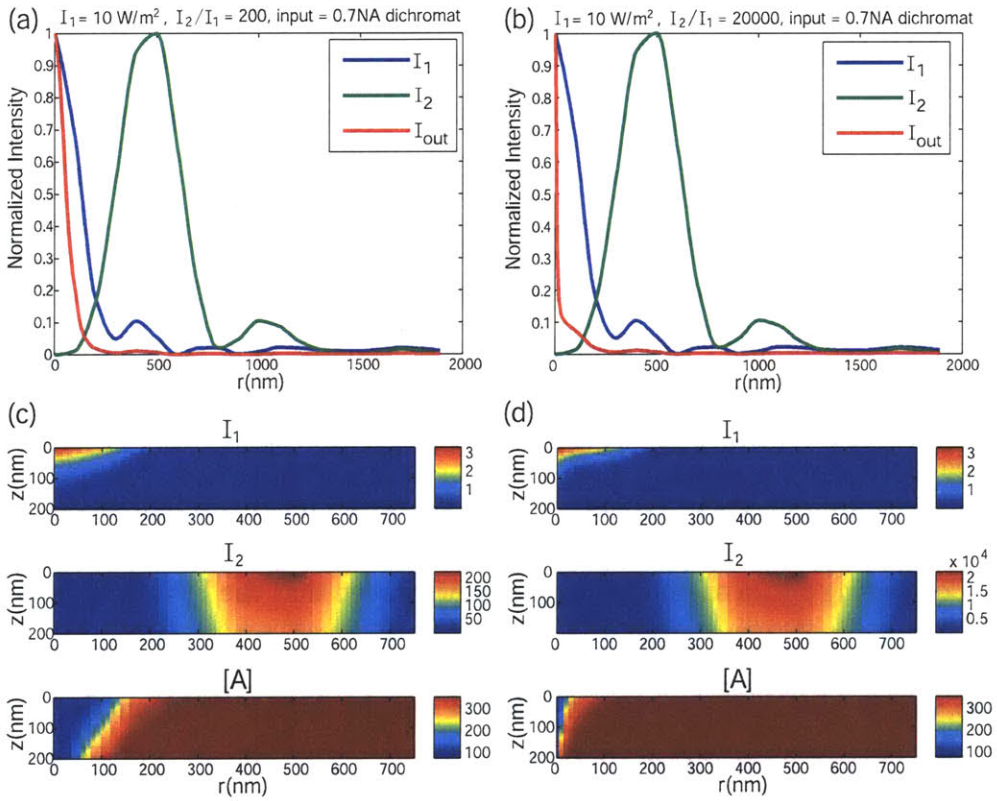

Figure 2-2: Typical input and output illuminations (upper figures) and cross-sectional profiles within the AML (lower figures) for two different incident I_2/I_1 ratios for NA = 0.7 dichromat input profiles. The left figures were obtained with $I_2/I_1 = 200$ and right figures with $I_2/I_1 = 20000$.

Figure 2-3 shows that Full-Width-Half-Maximum (FWHM) at the output plane of the AML scales down as a function of intensity ratios at steady state. The FWHM values were calculated for the λ_1 profile at the output AML plane with various incident I_2/I_1 ratios. The inputs profiles used were dichromat outputs with NA = 0.55, 0.7, and 0.83, illuminated with λ_1 and λ_2 plane waves. The radii of these dichromats are listed in Table 4.10, 4.11, and 4.13 in Chapter 4, respectively. The input illuminations of (e) and (f) are produced with a special dichromat design that only contains the first 115 zones of the NA = 0.83 dichromat design (160 zone). This special case was considered because in real dichromat fabrication, the zones beyond the 115th turned out to be unreliable or some turned detached during processing. The 115th zone of the NA = 0.83 design is a particularly wide zone (2.7 um), and so was easily identified and might have caused dose problems for neighboring zones due to proximity effect. The incident l_2/l_1 intensity ratios were adjusted by fixing $I_2 = 10$ W/m^2 and scaling I_1 down for figures on the left and by fixing $I_1 = 10$ W/m^2 and scaling I_2 up for figures on the right.

In general, FWHM decreases as I_2/I_1 ratio increases, except at very high I_2/I_1 ratios. Comparing figure on the left to the right, we see that the FWHM is compressed to smaller values when I_1 is fixed and I_2 scales up. For an ideal thermally stable AML, it was shown that the output FWHM should depend only on the ratio, not the absolute I_1 and I_2 intensity values [92]. However, when thermal instability exists, which is true for the azobenzene AML, higher I_1, I_2 intensities result in higher AML confinement because the stronger I_1, I_2 are, the less significant the thermal effect is.

Since thermal instability of the AML constantly drives the film back to its opaque state, this effect is analogous to adding a constant background of λ_2 illumination, which degrades the quality of the node at λ_2. To show how crucial the quality of the node is to the output FWHM, we conducted simulations by numerically adding different levels of λ_2 background noise to the NA = 0.7 dichromat input profiles, as shown in Fig. 2-4. Four levels of DC noise was added: 10%, 1%, 0.1%, and 0.01% with respect to the peak intensity of the λ_2 illumination. Figures 2-4(a)(c)(e)(g) shows the FWHM scaling curves at the four noise levels, and Figs. 2-4(b)(d)(f)(h)

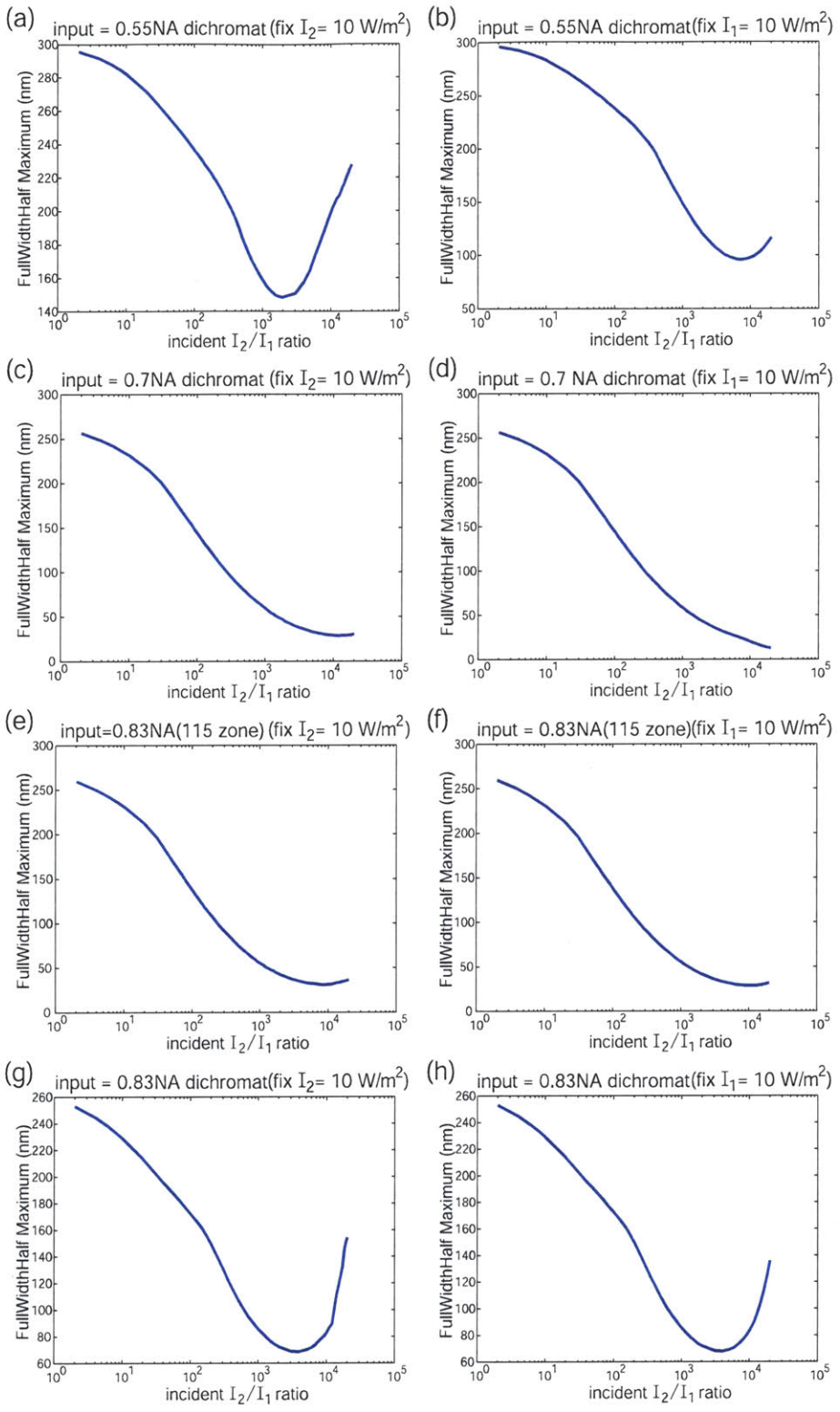

Figure 2-3: FWHM scaling with intensity ratios at steady state with input from dichromats with (a)(b) NA=0.55, (c)(d) NA = 0.7, (g)(h) NA = 0.83, and (e)(f) NA = 0.83 (115 zones). Figures on the left fix $I_2 = 10$ W/m^2 and figures on the right fix $I_1 = 10$ W/m^2 when changing incident I_2/I_1 ratios.

show the corresponding output profiles for $I_2/I_1 = 20000$ (the highest intensity ratio simulated for the FWHM scaling curve.) With lower noise, the intensity ratio where FWHM begins to increase is pushed higher and the minimum achievable FWHM is pushed smaller. As shown in Fig. 2-4(g), a noise level as low as 0.01% is required for the FWHM to not curve back up. Notice the unusual output beam profile in Fig. 2-4(h), consisting of a sharp spike and a broader base profile. This base profile grows with respect to the spike as the intensity ratio increases, and when the broad base reaches 50% of the peak intensity, the FWHM begins to increase. As will be seen in Chapter 4, the nodes of the NA = 0.55 and NA = 0.83 dichromats are not as deep as that of the NA = 0.7 dichromat. Combining the observations above explains why FWHM increases at high intensity ratios for NA = 0.55 and NA = 0.83 dichromats, as shown in Fig.2-3.

The effect of nonzero intensities at the node of the λ_2 beam is important because generating a perfect optical node is very difficult. This is especially true for node generation with diffractive optical elements (DOE). Since DOEs rely on canceling illuminations with opposite phase to achieve minimal illumination, the quality of the node is highly prone to fabrication errors. Any error in the placement or thickness of the DOE phase shifters will result in phase errors and nonzero intensities at the node.

In addition to the FWHM, peak-intensity efficiency of the absorbance modulation process is also important. We expect the absorbance modulation process to give rise to significant loss in energy because absorbance modulation works by absorbing light at locations where we don't want light to pass through. The efficiency that matters is the intensity efficiency at the point where light does pass through the AML. Figure 2-5 shows how the output/input λ_1 peak intensity efficiency changes with incident I_2/I_1 ratios. Similar to Fig. 2-3, the left figures are fixed at I_2=10W/m^2 and the right figures are fixed at I_1 = 10 W/m^2. We see that peak efficiency decreases with increasing I_2/I_1 ratio, *i.e.* increasing FWHM compression. This is also expected because when the size of the aperture in the AML enters the subwavelength regime, high spatial frequency components of the illumination decay evanescently. However, the decay is

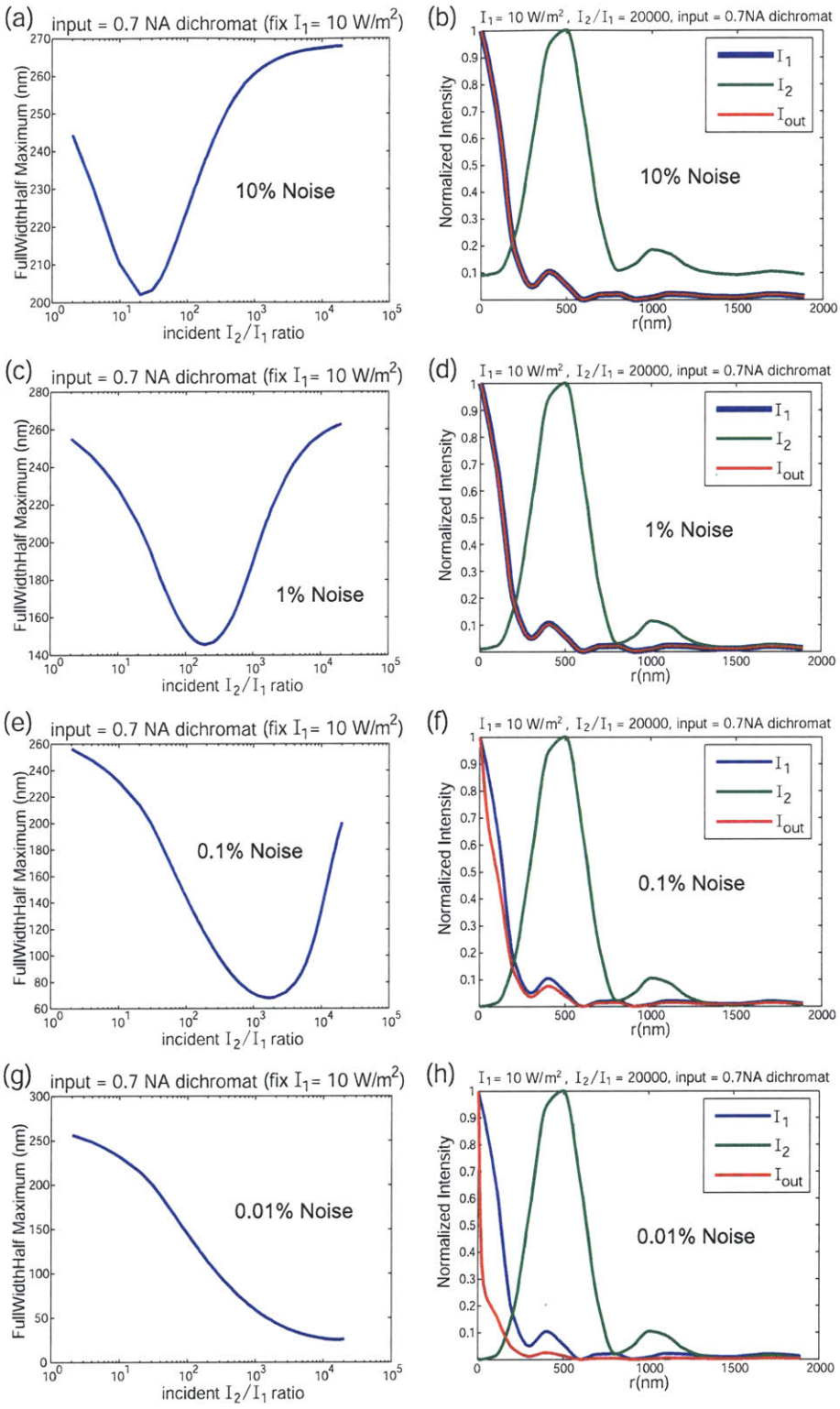

Figure 2-4: (a)(c)(e)(g) shows the FWHM scaling curves at the four noise levels, and (b)(d)(f)(h) shows the corresponding output profiles for $I_2/I_1 = 20000$

far more dramatic in the case of fixing I_2 and scaling I_1 down (left figures), giving efficiencies of only 10^{-6} at $I_2/I_1 = 20000$. This is probably the most significant difference between scaling I_2 up and scaling I_1 down and is again due to thermal instability of the azobenzene AML. The lower I_1 intensity is, the more significant thermal conversion of the AML to its opaque state is, and the more I_1 photons are absorbed during transmission.

Another kind of input profiles are generated by focusing a λ_1 plane wave and a λ_2 wave with a spiral phase front. The intensity profile of the focused round-shaped spot can be expressed analytically as a function of the NA of the focusing lens, as shown below.

$$I_1(\rho, z = 0) \propto J_1(\frac{2\pi NA\rho}{\lambda_1})/\rho \tag{2.11}$$

where J_1 is Bessel function of the first kind. A focused spot with spiral phase need to be computed numerically. The numeric form of a focused spiral-phase spot is

$$I_2(\rho, z = 0) \propto \frac{2\pi}{\rho^2} \int_0^{\frac{2\pi NA\rho}{\lambda_2}} J_1(\frac{2\pi NAx}{\lambda_2})xdx \tag{2.12}$$

Experimentally, spiral phase-fronts can be generated with a spiral-phase plate which imposes a phase equal to mθ, where m is an integer and θ is the azimuthal angle. The focused spiral-phase-plate (SPP) profiles and the corresponding focused spot profile for NA = 0.7, 0.83, and 0.95 are shown in Fig. 2-6. The focused SPP spots give rise to deep and wide nodes because the phase is ill-defined along the optical axis, so the intensity must go to zero. Figure 2-7(a) shows the FWHM scaling with I_2/I_1 ratios for SPPs with the three NAs. The same I_2/I_1 ratios as that in Fig. 2-3 to Fig. 2-5 were used for comparison. We see that FWHM decreases slower with SPP inputs than with dichromat inputs maybe because the doughnuts have flatter nodes, so the round-shaped spots don't feel as much compression until I_2/I_1 becomes large. Nevertheless, Fig. 2-7(b) shows that because the SPP nodes are assumed to be perfect, the FWHM curves does not curve back up at high intensity ratios. In addition to the monotonic decrease of the FWHM curves as I_2/I_1 ratios increase, Fig. 2-8 shows that peak efficiency is constant over I_2/I_1 and even the total energy

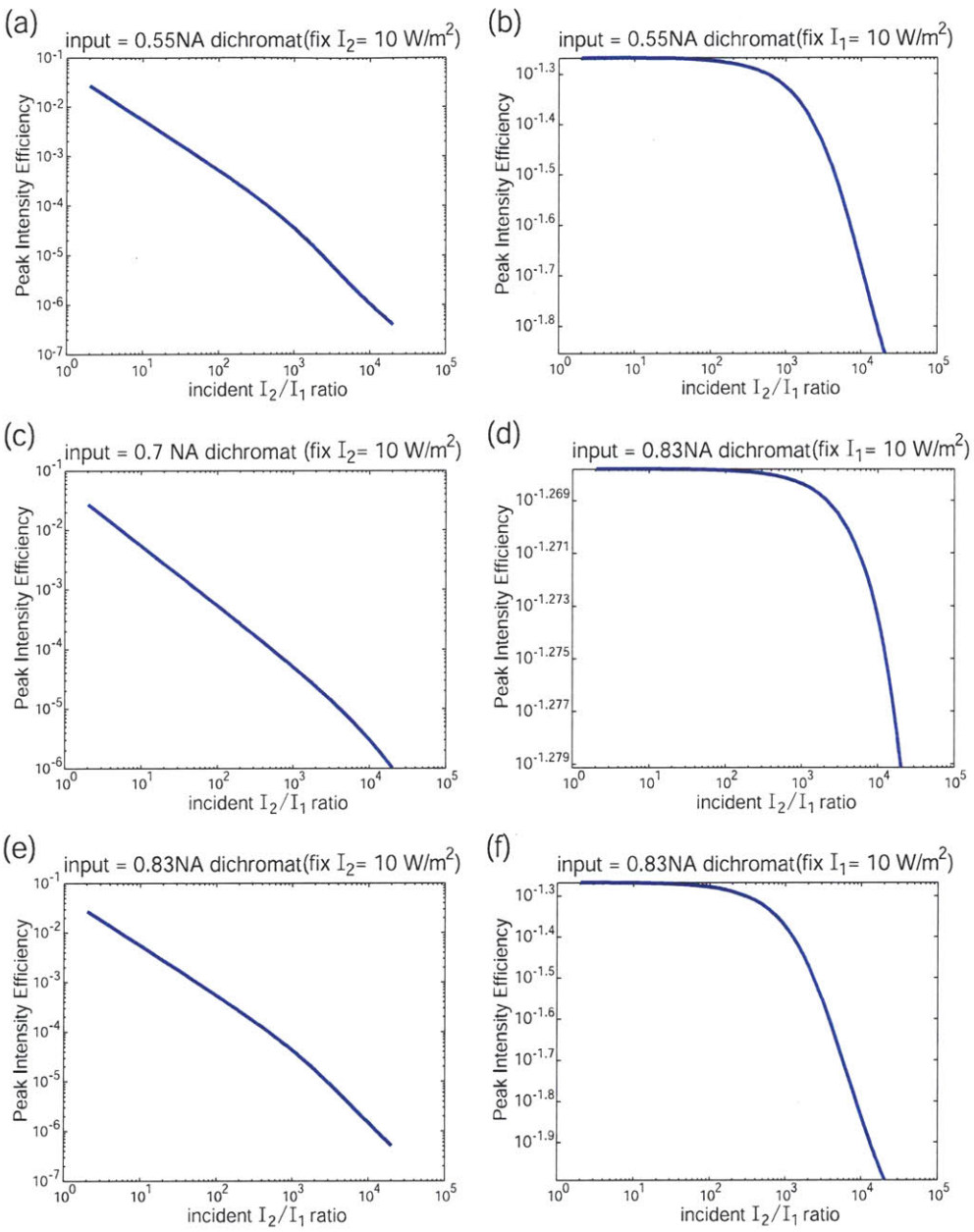

Figure 2-5: Peak intensity efficiency of the AML scales down with increasing I_2/I_1 intensity ratios at steady state. The input intensity profiles were the focal plane outputs of dichromats with (a)(b) NA = 0.55, (c)(d) NA = 0.7, and (e)(f) NA = 0.83. The incident I_2/I_1 ratios were adjusted by fixing $I_2 = 10$ W/m^2 and scaling I_1 down for figures on the left and by fixing $I_1 = 10$ W/m^2 and scaling I_2 up for figures on the right.

efficiency is quite reasonable, above 1%. In reality, SPPs are difficult to make perfect, so the quality of the node will still be limited by noise. Figure 2-9 shows simulations of FWHM scaling with various DC-noise levels in percentage of the λ_2 peak intensity. As shown in the figure, the higher the noise, the larger the minimum FWHM is and the faster FWHM starts bending up again when increasing I_2/I_1 ratio.

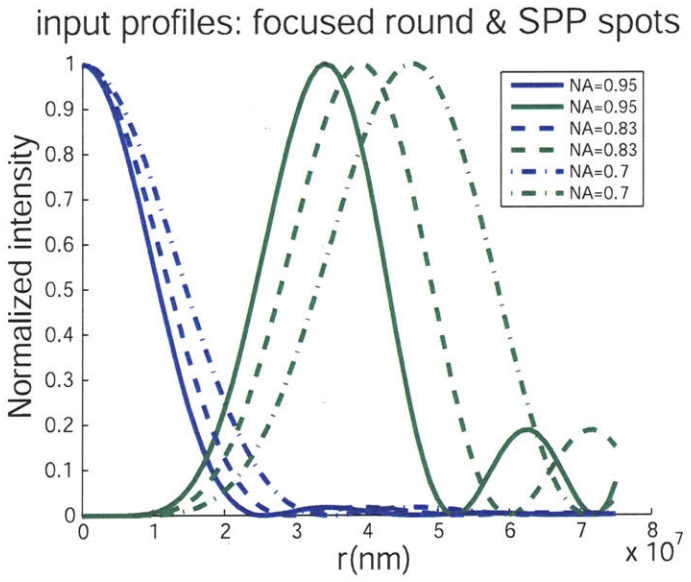

Figure 2-6: Focused spiral-phase-plate (SPP) profiles and the corresponding focused spot profiles for NA = 0.7, 0.83, and 0.95.

In summary, for thermally unstable AML, such as the azobenzene, higher absolute I_1 and I_2 intensities are desired to minimize background thermal conversions. Output FWHM far below 50 nm was predicted with input illuminations from an NA = 0.7 dichromat design. Peak intensity efficiencies with absorbance modulation is reasonable, ranging from 1-10% when I_1 is fixed at 10 W/m^2. Although the method is sensitive to noise in the l_2 node, some level of compression is still observed with reasonable noise levels below 10%. In practice, high intensity levels is always problematic due to difficulty in producing the light and concerns on overheating the substance that's illuminated. In Chapter 3, a thermally stable AML that works for $\lambda_1 = 325$ nm and $\lambda_2 = 633$ nm will be introduced. Steady-state simulations for that AML is also included as an option in the Matlab code in Appendix A.

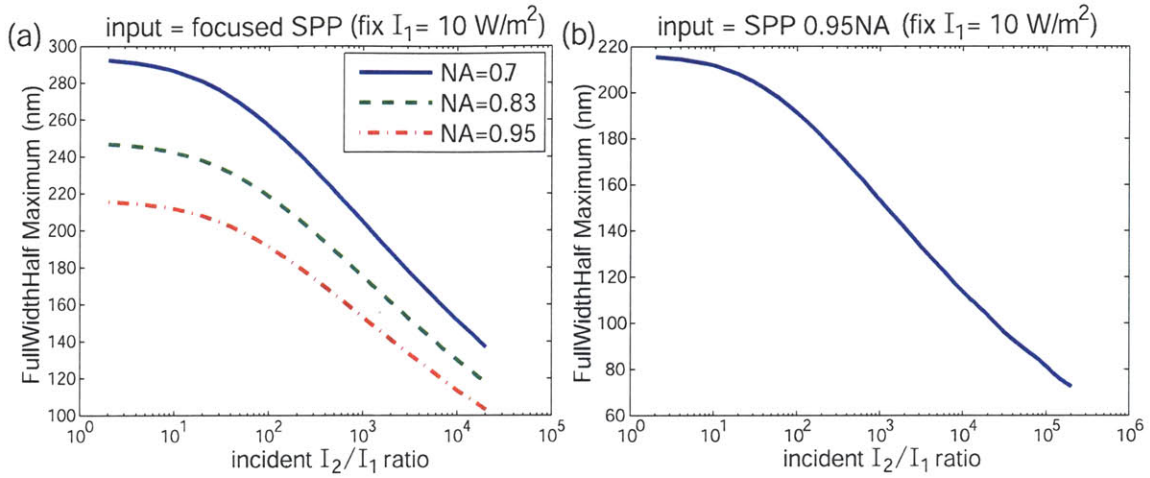

Figure 2-7: (a) FWHM scaling with I_2/I_1 ratios using illuminations shown in Fig. 2-6. (b) FWHM scaling for the NA=0.95 SPP to very high intensity ratios.

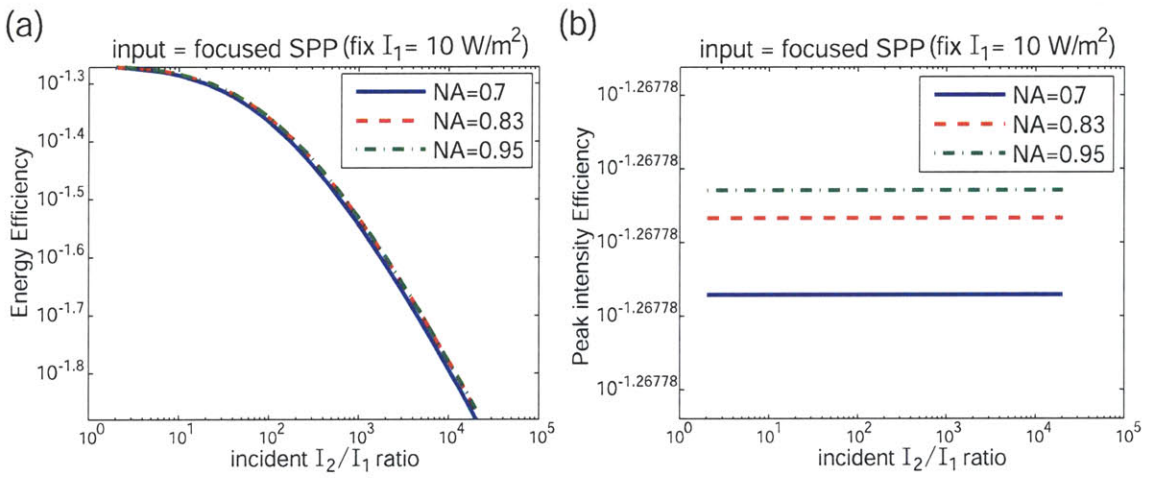

Figure 2-8: (a) Energy efficiency and (b) peak intensity efficiency of the AML as a function of I_2/I_1 intensity ratios at steady state. Energy efficiency is calculated by integrating the area under the input and output intensity profiles.

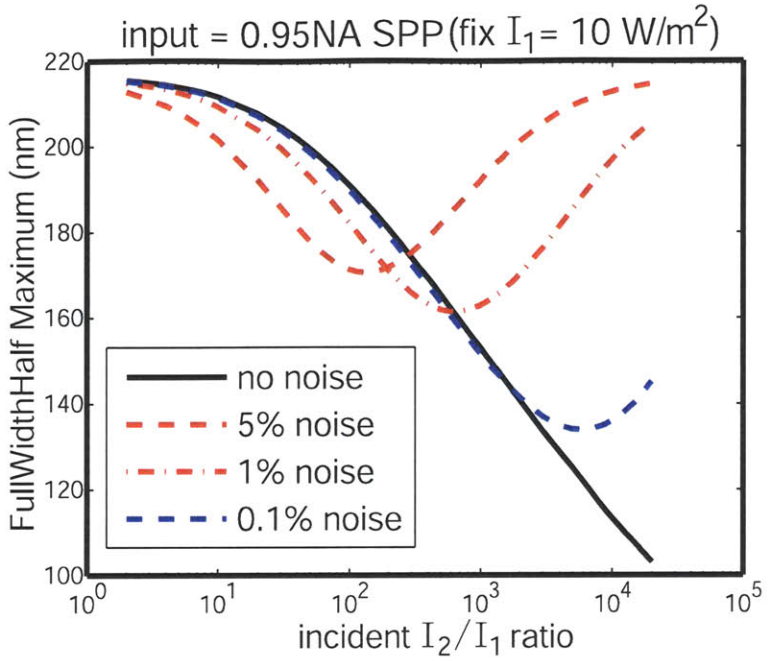

Figure 2-9: FWHM scaling as a function of I_2/I_1 ratio with different DC-noise levels in percentage of the λ_2 peak intensity.

2.4 Numerical Simulations with Finite Difference Method

2.4.1 The Transient Response

Although steady-state conditions apply most of the time, transient response can become significant when photo-kinetic transitions of the AML are slow or when I_1 is low and the AML is thermally unstable. A finite difference model was constructed to study the transient response of absorbance modulation systems. Transient response is important for estimating settling times of the AML in lithography and imaging applications. For lithography, transient exposures can be important when total exposure time is not much longer than settling time of the AML. For imaging, signals can be captured at different delays after illuminating the sample to capture or avoid the transient responses.

2.4.2 Distretization for Finite Difference

Finite-difference is a numerical method that approximates derivatives in differential equations as a finite difference. Finite difference is intuitive and often simple to implement, but can sometime lead to significant numerical error or instablilty if the step sizes or properties of the system of differential equations are not checked properly. When constructing finite difference analysis, the scheme of discretization can affect convergence and stability of the analysis. Most commonly used discretization schemes include explicit Euler, implicit Euler, and the Crank-Nicolson method. For example, $\frac{dX}{dt}|_{t_0}$ can be discretized in the following three ways:

$$\frac{dX}{dt}|_{t=t_0} \sim \frac{X(t_0+\Delta t)-X(t_0)}{\Delta t} \text{ (Implicit Euler)} \tag{2.13}$$

$$\sim \frac{X(t_0+1/2\Delta t)-X(t_0-1/2\Delta t)}{\Delta t} \text{ (Crank-Nicolson)} \tag{2.14}$$

$$\sim \frac{X(t_0)-X(t_0-\Delta t)}{\Delta t} \text{ (Explicit Euler)} \tag{2.15}$$

Depending on the differential equations to be solved, accuracy and condition for stability can be derived for each scheme. In general, explicit Euler is simplest to implement, but is usually subject to stability conditions and is only first order accurate, *i.e.* error is proportional to Δt. Implicit Euler is also first order accurate, but is unconditionally stable. The Crank-Nicolson discretization scheme was chosen to discretize Eq. 2.7, Eq. 2.8, and Eq. 2.9 because the scheme is second order accurate and unconditionally stable.

By rearranging the finite difference equations, the intensity values at the next z step (counted with the j index) and A-state concentration values at the next time step (counted with the n index) can be expressed as followed.

$$I_m^{j+1,n} = \frac{1-\frac{\Delta z}{2}[\epsilon_{1A}A^{j,n}+\epsilon_{1B}(A_0-A^{j,n})]}{1+\frac{\Delta z}{2}[\epsilon_{1A}A^{j,n}+\epsilon_{1B}(A_0-A^{j,n})]} I_m^{j,n}(\rho,z,t), m=1,2 \tag{2.16}$$

$$A^{j,n+1} = \frac{1-\frac{\Delta t}{2}[(\alpha_{1A}+\alpha_{1B})I_1+(\alpha_{2A}+\alpha_{2B})I_2+k_{BA}]}{1+\frac{\Delta t}{2}[(\alpha_{1A}+\alpha_{1B})I_1+(\alpha_{2A}+\alpha_{2B})I_2+k_{BA}]} A^{j,n}(\rho,z,t) + \frac{\Delta t(\alpha_{1B}I_1+\alpha_{2B}I_2+k_{BA})}{1+\frac{\Delta t}{2}[(\alpha_{1A}+\alpha_{1B})I_1+(\alpha_{2A}+\alpha_{2B})I_2+k_{BA}]} A_0 \tag{2.17}$$

The system of equations are solved in a two-step iteration manner, first in z and then in t. The illumination profiles at the input plane of the AML gives the initial conditions in z, $I_1(\rho, z = 0, t)$ and $I_2(\rho, z = 0, t)$, at all times. At the first time step, the initial conditions are first iterated over the entire z range with Eq. 2.16, assuming uniform AML absorbance $[A](\rho, z, t = 0) = A_0$. After the first iteration in z, Eq. 2.17 is applied to evaluate $[A](\rho; z)$ for the next time step. After each time step, $I_1(\rho; z; t)$ and $I_2(\rho; z; t)$ are re-iterated from $z = 0$ to $z =$ (AML thickness) with the updated $[A]$ values. Updated $I_1(\rho; z; t)$ and $I_2(\rho; z; t)$ values are then used to step $[A]$ forward to the next time step, and so on so forth.

Snap shots of the transient response of an AML illuminated by an NA = 0.7 dichromat input profiles with $I_2(\text{peak})/I_1(\text{peak}) = 200$ (fixed I_1 at 10 W/m^2) is shown in Fig. 2-10. Steady state was reached in about 100 ms. As the I_1 profiles penetrate gradually into the AML, the A-state (opaque form) concentration decreases and the aperture for λ_1 illumination is formed. Note that because of the high λ_2 intensity, I_2 profiles penetrate the AML much faster and cause A-state concentration to drop at where λ_2 illuminates first. The fact that A-state concentration also drops upon λ_2 illumination is because Φ_{2AB} is not zero, which means that λ_2 illumination also converts the AML from A to B. Therefore the contrast between Φ_{2AB} and Φ_{2BA} is also important to the final contrast of the aperture in the AML. The steady-state $[A]$ value outside of the aperture region is consistent with that calculated in Fig. 2-2 using the steady-state model. The colors appear differently in the two figures only because the color bar was auto-scaled in Fig. 2-2, but was fixed at maximum $[A] = A_0$ in Fig. 2-10.

In order to look at the transient response more quantitatively, the transient response of the concentration of A-state (opaque from) on the optical axis, *i.e.* $[A](\rho = 0, z =$ AML thickness$, t)$, is shown in Fig. 2-11 with NA = 0.7 dichromat inputs at various input intensities. Intuitively, lower intensity gives rise to slower transient responses because photo-switching rates are proportional to intensity. Figure 2-11 shows that the settling time significantly lengthen by lowering I_1 input intensity. Figure 2-11(a) and (b) were simulated with $I_2 = 10$ W/m^2 and an I_2/I_1

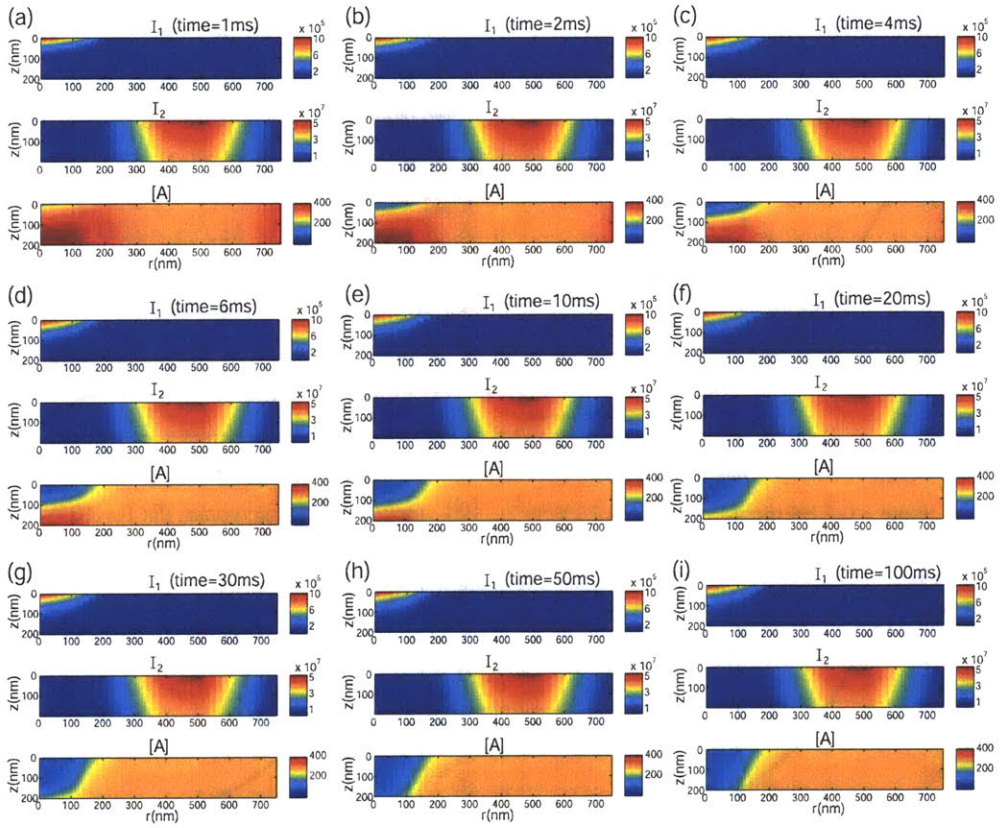

Figure 2-10: Snap shots of the transient response of an AML illuminated by NA=0.7 dichromat input profiles with $I_2/I_1 = 200$, fixing $I_1 = 10\text{W/m}^2$.

ratio of 2 and 200, respectively. Because λ_2 intensity on the optical axis is small, this effect is not caused of the much higher I_2/I_1 ratio. Figure 2-11(c) and (d) were simulated with $I_1 = 10$ W/m^2 and an I_2/I_1 ratio of 2 and 20000, respectively. The two figures show no obvious difference in settling time at $\rho = 0$. Higher I_2 intensity also leads to faster transient responses. Figure 2-12(a) and (b) shows the transient response of $[A]$ in the same simulation as Fig. 2-11(c) and (d), but at 90 nm away from the optical axis, *i.e.* $\rho = 90$ nm. The settling time is significantly reduced with higher I_2 intensity.

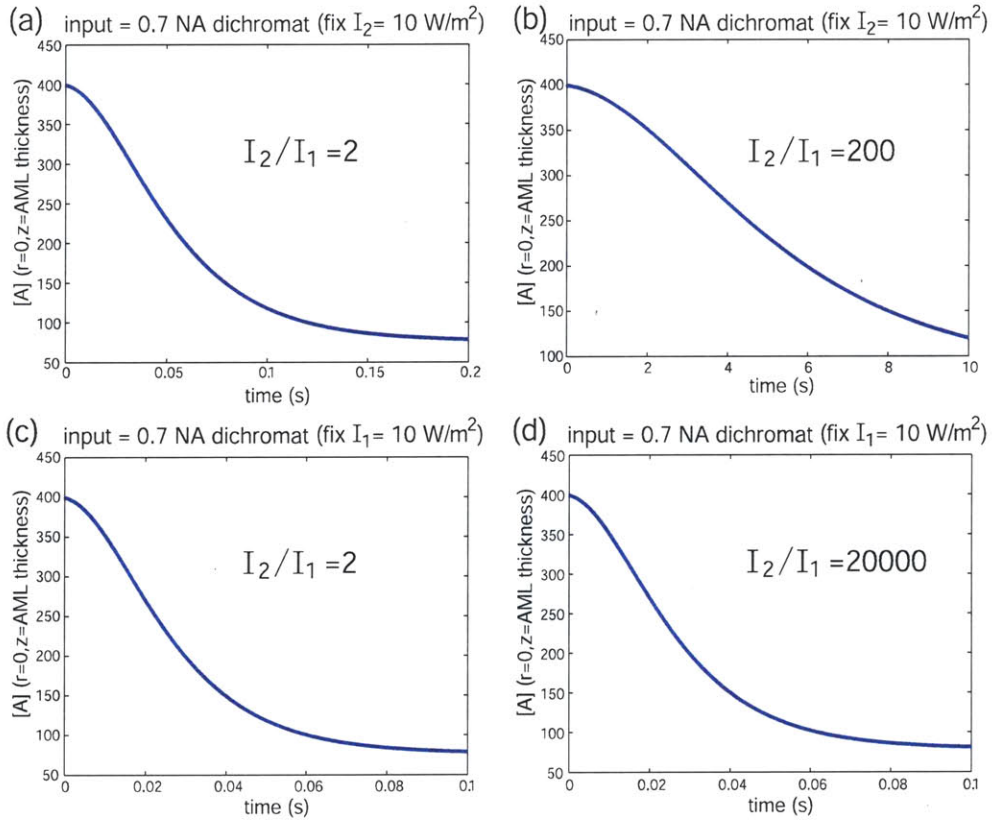

Figure 2-11: Transient response of the concentration of A-state (opaque from) on the optical axis, $[A](\rho = 0, z = \text{AML thickness}, t)$, with NA = 0.7 dichromat inputs. Settling time significantly lengthen by lowering I_1 input intensity.

Another way to quantify the transient response is to monitor FWHM change over time at the AML output plane. Figure 2-13 shows this transient response with various I_2/I_1 ratios at a fixed λ_1 illumination intensity of 10 W/m^2. Figure 2-13(b) shows that when the I_2/I_1 ratio is reasonably high, the FWHM decreases monotonically

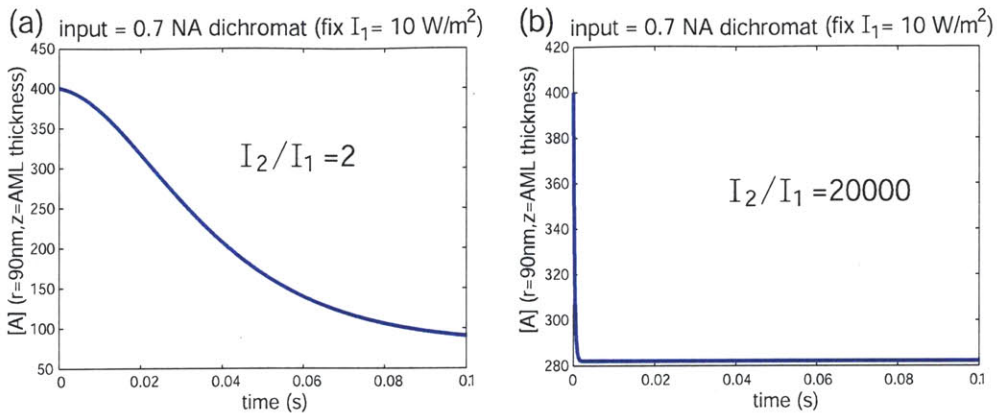

Figure 2-12: Transient response of the concentration of A-state (opaque from) 90nm away from the optical axis, $[A](\rho = 90 \text{ nm}, z = \text{AML thickness}, t)$, with NA = 0.7 dichromat inputs. Settling time significantly shorten with high I_2 intensity.

with time as the aperture in the AML is formed. However, when I_2/I_1 is low, the output FWHM first decreases and then rises to its steady-state values, as shown in Fig. 2-13(a). Although the steady-state FWHM is quite large, transient FWHM can be much smaller when only the peak of I_2 penetrates through the AML. This is very similar to the concept of contrast enhancement layer (CEL) for lithography [94], which utilizes transient response of a photo-bleachable film to enhance contrast.

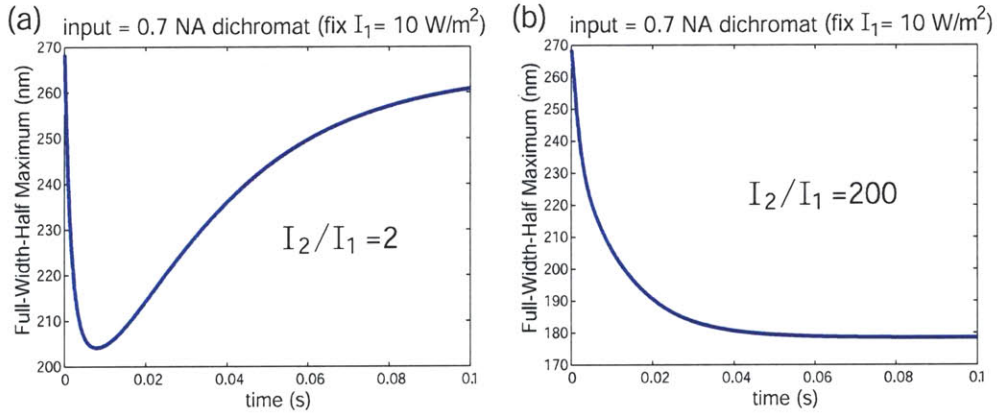

Figure 2-13: Transient FWHM at AML output plane for $I_1 = 10\text{W/m}^2$ with I_2/I_1 ratios of 2 and 200.

In the next chapter, we will discuss future work on applying the AML as a reversible CEL. Figure 2-14 shows snap shots of the transient response with only λ_1

input from an NA = 0.7 dichromat ($I_1 = 10$ W/m^2) to emulate the use of the azobenzene AML as a CEL. Because A-state concentration decreases faster with higher λ_1 illumination, a transient aperture in the AML is formed at the peaks of the λ_1 profile. Background profiles outside of the peaks are therefore suppressed, so the contrast is enhanced. One thing to note is that unlike cases with the λ_2 doughnut-shaped illuminations, contrast enhancement without λ_2 only exist in transient. In steady state, as shown in the bottom plot of Fig. 2-15(a), the AML is converted to the same steady-state $[A]$ value everywhere λ_1 illumination exists, so the AML does not provide additional contrast. Figure 2-15 shows a summary of I_1, I_2, $[A]$ profiles after 100 ms of iteration and the steady-state $[A]$ profiles for $I_2/I_1 = 0$, 2, 200, 20000 with I_1 fixed at 10 W/m^2 using NA = 0.7 dichromat input profiles. The steady-state solution here is approximated by applying steady-state conditions at the end of the 100 ms iteration. In steady state, Eq. 2.9 can be simplified to directly associate $[A]$ at each point with its local I_1, I_2 intensities as followed.

$$A_{steadystate} = \frac{\alpha_{1B} I_1 + \alpha_{2B} I_2 + k_{BA}}{(\alpha_{1A} + \alpha_{1B}) I_1 + (\alpha_{2A} + \alpha_{2B}) I_2 + k_{BA}} A_0 \tag{2.18}$$

The same result can also be derived from Eq.2.17 by setting $A^{n+1}_{matrix} = A^n_{matrix}$.

$$\begin{aligned} A^{n+1}_{matrix} = & \frac{1 - \frac{\Delta t}{2}[(\alpha_{1A} + \alpha_{1B}) I_1 + (\alpha_{2A} + \alpha_{2B}) I_2 + k_{BA}]}{1 + \frac{\Delta t}{2}[(\alpha_{1A} + \alpha_{1B}) I_1 + (\alpha_{2A} + \alpha_{2B}) I_2 + k_{BA}]} A^n_{matrix} \\ & + \frac{\Delta t(\alpha_{1B} I_1 + \alpha_{2B} I_2 + k_{BA})}{1 + \frac{\Delta t}{2}[(\alpha_{1A} + \alpha_{1B}) I_1 + (\alpha_{2A} + \alpha_{2B}) I_2 + k_{BA}]} A_0 \end{aligned} \tag{2.19}$$

By plugging in this steady state expression at the end of the iteration, the solution is forced to jump close to steady state much faster. The "steady state" solution shown in Fig. 2-15 was computed by applying only one steady state iteration for $[A]$. All simulations in the section was conducted with time step of 0.1 ms, depth (z) step of 1 nm, and radial (r) step of 10 nm.

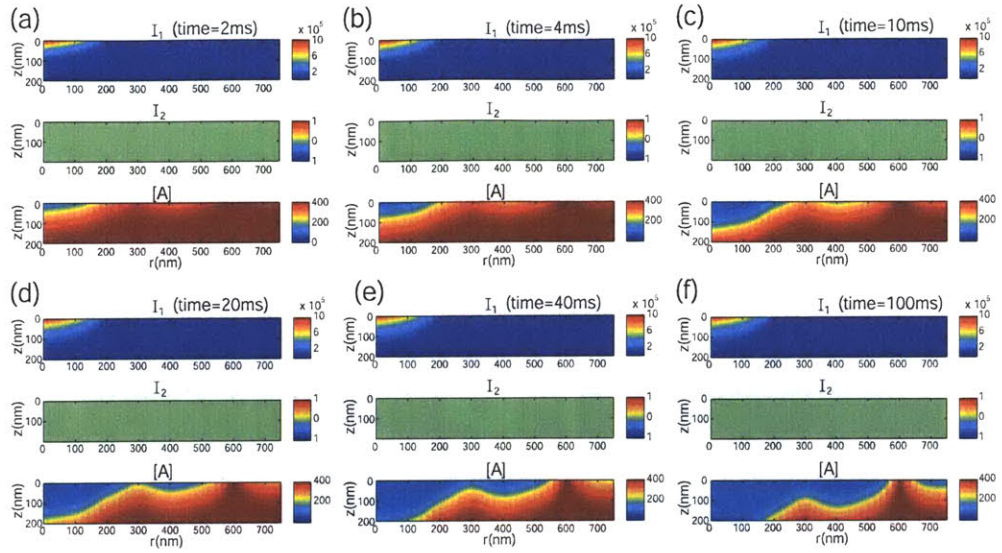

Figure 2-14: Snap shots of the transient response of an AML illuminated by NA=0.7 dichromat input profiles with $I_2/I_1 = 0$ (no λ_2 illumination), fixing $I_1 = 10\text{W/m}^2$. The radial axis, ρ, is denoted as "r" in the plot. $\rho = 0$ corresponds to the optical axis of the dichromat.

2.4.3 Convergence of the Finite Difference Simulations

The convergence of the finite difference schemes is defined as the difference between the iteration for the next step and the exact value at that step. The exact values can be obtained in theory either by setting the step sizes to zero or by expanding finite difference equations with infinite seies. The convergence is often determined by computing the Taylor series expansion of the exact function and subtracting the finite difference expression with the series.

For example, the error for Eq. 2.16 for time $t = 0$ $(n = 0)$ is

$$\begin{aligned}
error &= I^{j+1,0} - I((j+1)dz, 0) \\
&= \frac{1 - \frac{\Delta z}{2}\epsilon_A A_0}{1 + \frac{\Delta z}{2}\epsilon_A A_0} I^{j,0} - (I^{j,0} + \frac{\partial I}{\partial z}\Delta z + \frac{1}{2}\frac{\partial^2 I}{\partial z^2}\Delta z^2 + \dots) \\
&\sim \frac{I^{j,0} - \frac{\Delta z}{2}\epsilon_A A_0 I^{j,0} - I^{j,0} - \frac{\Delta z}{2}\epsilon_A A_0 I^{j,0} - \frac{\partial I}{\partial z}\Delta z - \frac{\Delta z^2}{2}\epsilon_A A_0 \frac{\partial I}{\partial z} - \frac{1}{2}\frac{\partial^2 I}{\partial z^2}\Delta z^2}{1 + \frac{\Delta z}{2}\epsilon_A A_0} \\
&= -\frac{\epsilon_A A_0 \frac{\partial I}{\partial z} + \frac{\partial^2 I}{\partial z^2}}{2(1 + \frac{\Delta z}{2}\epsilon_A A_0)}\Delta z^2, \text{ because } \frac{\partial I}{\partial z} = -\epsilon_A A_0 I^{j,0} \text{ from Eq. 2.7 and 2.8} \\
&= O(\Delta z^2), \text{ when } \Delta z \epsilon_A A_0 << 1 \qquad (2.20)
\end{aligned}$$

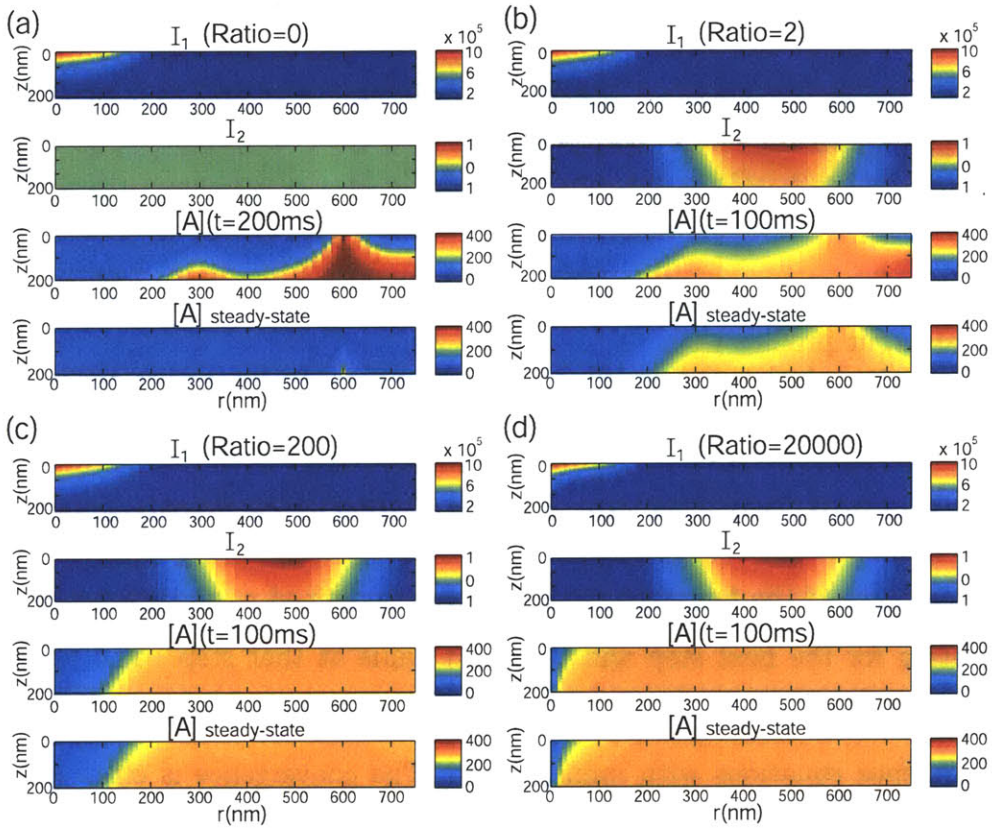

Figure 2-15: Summary of I_1, I_2, $[A]$ profiles after 100 ms of iteration and the steady-state $[A]$ profiles for $I_2/I_1 = 0$, 2, 200, 20000 with I_1 fixed at 10 W/m^2 using NA = 0.7 dichromat input profiles. The radial axis, ρ, is denoted as "r" in the plot. $\rho = 0$ corresponds to the optical axis of the dichromat.

The final convergence errors for Eq. 2.16 and Eq. 2.17 are shown below.

$$I^{j+1,n} - I((j+1)dz, ndt) = -\frac{c_1\frac{\partial I}{\partial z} + \frac{\partial^2 I}{\partial z^2}}{2(1+\frac{\Delta z}{2}c_1)}\Delta z^2 = O(\Delta z^2), \text{ when } c_1\Delta z << 1 \quad (2.21)$$

$$A^{j,n+1} - A(jdz, (n+1)dt) = -\frac{c_2\frac{\partial A}{\partial t} + \frac{\partial^2 A}{\partial t^2}}{2+c_2\Delta t}\Delta t^2 = O(\Delta z^2), \text{ when } c_2\Delta t << 1 \quad (2.22)$$

where $c_1 = \epsilon_{mA}A^{j,n} + \epsilon_{mB}(A_0 - A^{j,n})$ ($m = 1$ or 2), and $c_2 = I_1^{j,n}(\alpha_{1A} + \alpha_{1B}) + I_2^{j,n}(\alpha_{2A} + \alpha_{2B}) + k_{BA}$.

Convergence plots are generated below by comparing values of I_1, I_2, and $[A]$ at the final iteration time, 100 ms, to the values obtained with the finest grid size (in Δz or Δt). Figure 2-16 shows the convergence plots for the scheme with a test input of NA = 0.7 dichromat profiles, $I_1 = 10$ W/m^2, and $I_2/I_1 = 2000$. The L_1 norm is defined as the sum of the magnitude of the errors.

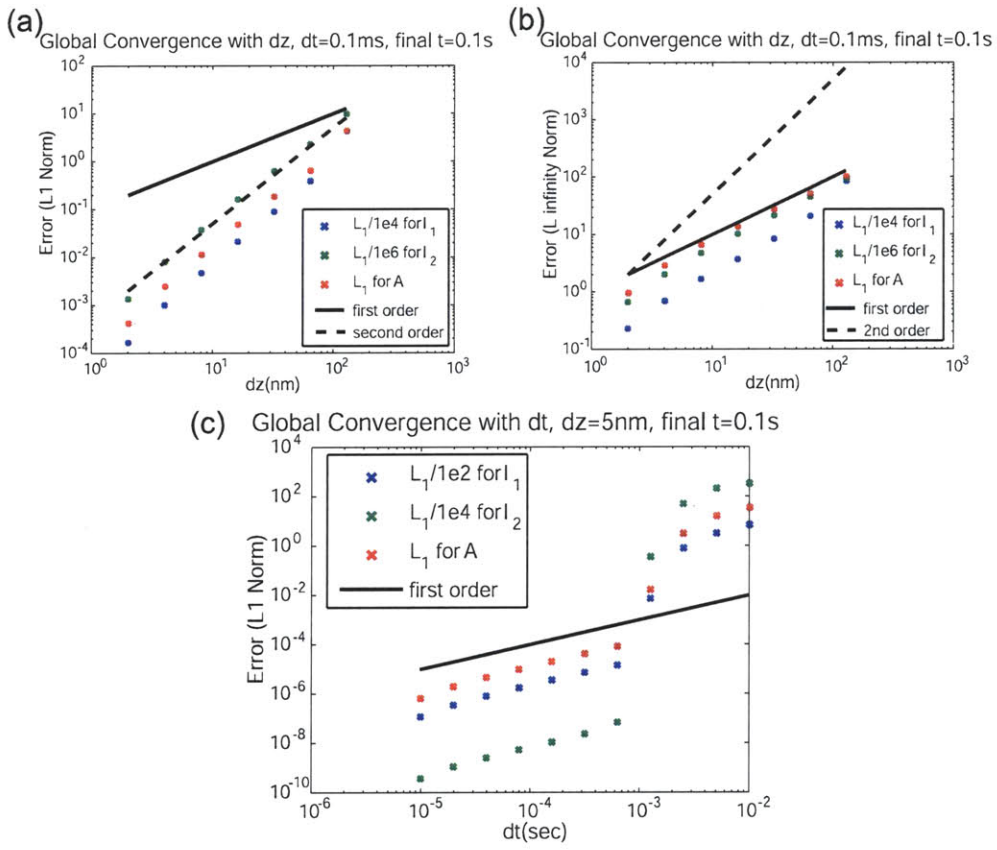

Figure 2-16: Convergence plot for the Finite Difference model with a test input of NA=0.7 dichromat profiles, $I_1 = 10$ W/m^2, and $I_2/I_1 = 2000$.

As seen in Fig. 2-16, the L_1 norm shows second order convergence with Δz, but only first order with dt. Because I_1, I_2, and $[A]$ all use Crank-Nicolson decretization for z and for t, second order convergence was expected for both. The reason for only first order convergence in time could be because that the scheme is second order accurate for one step, but error accumulates over the total number of steps to reach the final z or time. The number of steps is inversely proportional to Δz or Δt, so the total convergence becomes only first order. The convergence with dz is second order because not only I_1 and I_2 at the output plane (z = AML thickness) is considered. That's also why in Fig. 2-16(b), the L_∞ norm for the errors in z, defined as the maximum error among all samples, also shows first order convergence. Another thing to note is the inequality assumption made in the derivation of Eq. 2.16 and Eq. 2.17. In Eq. 2.16, if we use the maximum Δz = AML thickness = 200 nm, the largest value of $\Delta z \epsilon_{1A} A_0$ is 24. In order to satisfy the requirement for second order convergence, $\Delta z \epsilon_{1A} A_0 << 1$, it is required that $\Delta z << 10$ nm, which holds for half of the convergence curve with Δz. In contrast, for Eq. 2.17, c_2 is input dependant and can become quite large easily. In this case, α 10^{-5}-10^{-4} and I_1, I_2 $\sim 10^6$, so c_2 is on the order of 10^2. As a result, $c_2 \Delta t$ term in the denominator becomes non-negligible when Δt is larger than 10^{-3}. The scheme then becomes only first order accurate for each step and causes the error jump in Fig. 2-16(c). The time step used for most simulations was 10^{-4}, which falls in the left portion of the curve.

2.5 Future Work

2.5.1 Full Photon Transport Description

With direct discretization of incident angle θ of the incoming photon, the same equations and discretization schemes used in this thesis can be extended to solve the more general photon transport equation with photons transporting in both z and ρ. The θ distribution of the incident beam pattern can be approximated to be proportional to amplitude square of the Fourier Transform of the incident E-field amplitude and

phase as shown below. Instead of assuming photons travel only in the z-direction, *i.e.*,

$$L(r, \hat{\Omega}, t) = I(r, t)\delta(\hat{\Omega} - \hat{z}) \tag{2.23}$$

the radiance L now has an angular spread at the top surface of the AML:

$$L(r, \hat{\Omega}, t) = I(r, t)D(\hat{\Omega}) = I(r, t)D(x', y') \tag{2.24}$$

$$D(x', y') = \frac{|FT\{E(x'y')\}|^2}{\int_{y'}\int_{x'}|FT\{E(x'y')\}|^2 dx'y'} \tag{2.25}$$

This angular distribution will now depend on the numerical aperture of the incident beam, proving a more realistic initial condition for photon transport. Note that this extension still ignores scattering, *i.e.*, photons don't change transport directions after propagating into the AML. As a result, photons propagating at each angle are independent and can be solved separately. Considering this angle spread is also not the same as taking into account diffraction. The complexity of the problem will simply be $O(n_\rho \times n_\theta)$, where n stand for the number of discretization. Taking into account scattering or diffraction effects within the AML would be much more complicated as the formulation will no longer be decoupled and equations will become even more non-linear.

2.5.2 Consider Important Physical Phenomena: Diffraction

In all models we have considered so far, diffraction or coherence of the illumination source was not considered. With the sub-wavelength length scales of the aperture size in the AML and thickness of the AML, the effect of diffraction is not negligible. As diffraction normally smoothens the distribution of light and attenuates high spatial frequency components, the output PSF profiles obtained without considering diffraction are likely optimistic, predicting tighter light confinements and higher output efficiencies. Although other physical models that take diffraction into account can be applied, few simplifying assumptions can be applied at the scale of this problem. For example, the beam propagation method (BPM) takes into account diffraction and

has a very similar layer by layer numerical implementation to the photon transport model constructed in this work. However, since BPM solves the Helmholtz equation only under slowly varying envelope approximation, which requires long propagation length compared to the wavelength of light and slow variation of material properties along the propagation direction, it is unclear whether a BPM implementation can outperform the photon transport model. The most promising, but also most computationally expensive solution is the finite-difference time-domain (FDTD) method.

An example of a FDTD simulation conducted for this work to account for PSF broadening due to diffraction along propagation is shown below. In order to illustrate the effect of the thickness of the PVA layer on the linewidth of the pattern, Dr. Rajesh Menon simulated the transmission of light through a subwavelength aperture in a metal film using custom software that implements the FDTD method [95]. When a subwavelength aperture is illuminated, evanescent high-spatial frequency components of the transmitted light decay exponentially away from the aperture, increasing the FWHM of the transmitted light. The illumination of a one-dimensional aperture with a width of 75 nm was simulated with a plane wave with a wavelength of 325 nm, as shown in the schematic in Fig. 2-17(a). The electric field of the incident wave was polarized normal to the plane of the figure. The time-averaged intensity of the scattered light was calculated at steady state. Cross sections of the normalized intensity distribution in planes parallel to the aperture at varying distances from the aperture were computed and plotted in Fig. 2-17(b). Clearly, the transmitted light is substantially broadened with distance from the aperture. Furthermore, the peak intensity at the center of the line also falls exponentially with distance from the aperture.

In order to apply this FDTD simulation for simulating the optical fields when photons transport through the AML, we need to define an AML material with corresponding photo-dynamic properties. This is challenging for most FDTD solvers because most solvers assume static material profiles throughout the time period of simulation. In fact, as mentioned in section 2.2.2, the time scale required for the FDTD transport solution to reach steady state is much shorter than the photo-switching

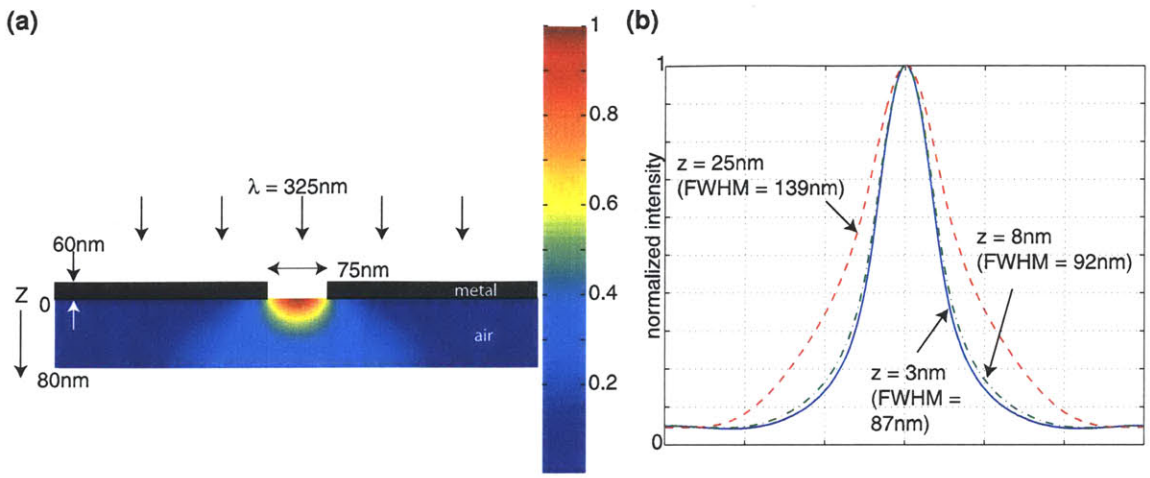

Figure 2-17: Effect of propagation distance from a line aperture on the FWHM of the line spread function. (a) FDTD simulation of light transmission through a one-dimensional subwavelength aperture. The incident light was assumed to be a uniform plane wave with an electric field polarized into the plane of the figure. The metal was assumed to have $\epsilon = 0.85 + i2.01$. The metal film was placed in air. (b) Cross sections of the normalized intensity distributions at different distances from the aperture. The evanescent high spatial frequencies die away from the aperture, and the linewidth (defined by the FWHM) increases.

speed of the AML. Therefore, approximating the AML as a static layer over the time period of the FDTD simulation is a legitimate assumption. One can use the intensity profiles obtained with the steady-state FDTD transport simulation to compute the corresponding steady-state $[A]$ profile (which is not the real steady-state $[A]$ because the intensity profiles has not yet responded to the change of $[A]$) and then conduct FDTD transport simulation again with the new $[A]$ profile in the AML. This iterative process can be carried out until solutions converge and the true steady-state solutions of I_1, I_2, and $[A]$ are obtained for the given incident light intensities. This simulation process can also be viewed as simply replacing Eq. 2.7 and Eq. 2.8, the linear transport equations, with an FDTD solver to properly propagate light through the AML. As a result, obtaining transient response with this process is also possible by choosing a proper time step for Eq. 2.9 and step forward the concentration profile $[A]$ using finite difference method. During each iteration, intensity distributions are updated with the steady-state FDTD transport solution using the updated $[A]$

profile. Nevertheless, as FDTD simulations are often time consuming, the number of iterations required to obtain full FDTD description of the photon transport through the AML can quickly exceed practical limits. Tradeoffs among size of the time step in Eq. 2.9, FDTD grid size, and computational resources (time and memory) would be required.

Chapter 3

One-Dimensional Demonstration of Absorbance Modulation with Interference Lithography

3.1 Previous Work on Absorbance Modulation Lithography

After proposal of the concept of absorbance modulation in 2005 [92], we conducted a series of experiments to prove the concept. We first reported confinement of a uniform beam of light ($\lambda_1 = 405$ nm) at the nodes of a standing wave ($\lambda_2 = 532$ nm) via absorbance modulation[96]. In this set of experiment, a thin polymer film containing a photochromic azobenzene side chain, referred to as the absorbance modulation layer (AML,) was exposed to a standing wave at λ_2 and a uniform beam at λ_1, resulting in alternate regions of high and low absorbance. A schematic of the optics setup is shown in Fig. 3-1. Using photoresist exposures, we mapped out the λ_1 light intensity distribution localized in low-absorbance regions and compared that with our theoretical model. Continuous lines with width as thin as 100nm ($\lambda_1/4$) were recorded in photoresists, as shown in Fig. 3-2. Although theory predicts that the line width of that λ_1 distribution is primarily determined only by the intensity ratio at the

two wavelengths and can potentially be confined to scales far below the diffraction limit, this prediction requires the AML be thermally stable and the photoresist be completely insensitive to λ_2. In experiments with the azobenzene AML, both criteria were not met as the azobenzene AML is thermally unstable and the photoresits used are not completely insensitive to λ_2=523 nm illumination.

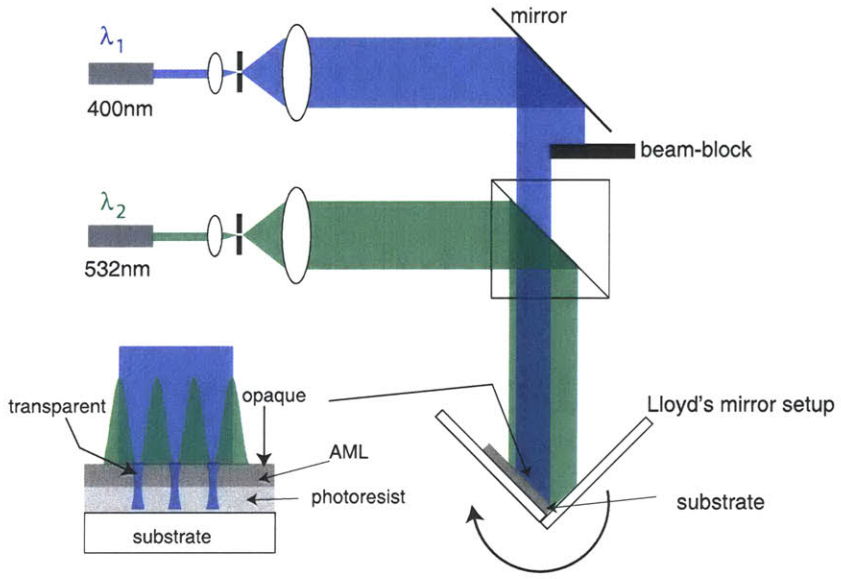

Figure 3-1: Schematic of Lloyd's mirror setup for AMOL. The standing-wave illumination at λ_2 creates local apertures for λ_1, through which the underlying photoresist is exposed.

For optimum performance, it is essential that the photochromic molecules are thermally stable; otherwise, the size of the writing beam becomes dependent on the absolute intensities rather than their ratio alone, thermal instability also limits scale down. If the photochromic molecule in the transparent state is thermally unstable, then at low λ_1 intensities, thermal back-reaction can overwhelm the forward reaction and close the aperture. Figure 3-3 shows a comparison of the simulated full-width-at-half-maximum (FWHM) of the transmitted light at λ_1 as a function of the ratio of intensities at λ_2 over λ_1 with and without thermal instability. The photochromic parameters of 1,2-bis(5,5'-dimethyl-2,2'-bithiophen-yl) perfluorocyclopent-1-ene (compound 1), the AML which will be described in more detail later, were assumed. The

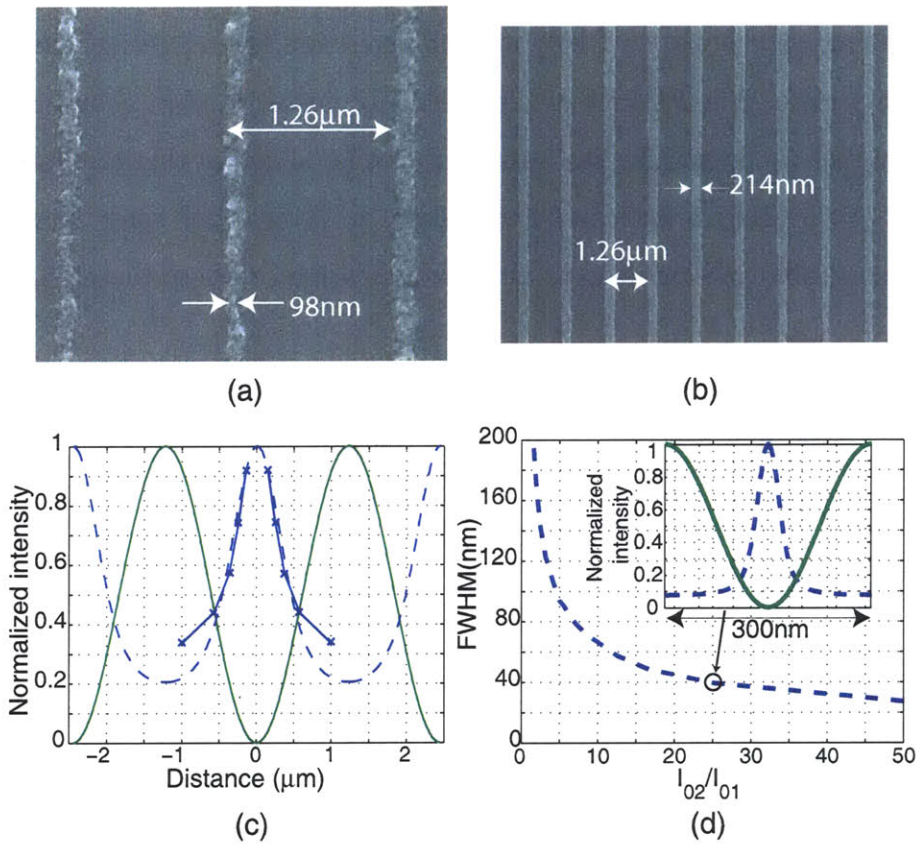

Figure 3-2: (a),(b) Scanning electron micrographs of lines exposed in negative photoresist. The average linewidth of the middle line in Fig. 3-2(a) is about 117 nm. The period of the lines is determined by λ_2 and the angle in the Lloyd's mirror setup. (c) Light intensity distribution of the localized fields underneath the AML for a standing-wave period of 2.45 μm. The simulated data is shown with dashed lines. The experimental data is overlaid (solid line) for comparison. (d) Calculated FWHM as a function of I_{02}/I_{01} for a standing-wave period of 300 nm. The parameters of the AML listed in Fig. 3-4 were used in all the simulations.[96]

AML is thermally stable and its FWHM curve is shown by the solid line in Fig. 3-3. An artificial thermal rate constant of $5 \times 10^{-4} s^{-1}$ was assumed for the dashed curve. The incident illumination is modeled as standing waves with periods of 350 nm (λ_2 = 633 nm) and 170 nm (λ_1 = 325 nm). Both curves were calculated by decreasing the peak intensity of the λ_1 standing wave while maintaining the peak intensity of the λ_2 standing wave equal to 1 kW/m^2 and repeating the numerical simulation for each intensity ratio. The FWHM of the thermally unstable AML shows a minimum. Although higher λ_1 intensities can result in smaller sizes for this minimum, the intensity at λ_2 will need to be scaled accordingly to achieve the same intensity ratio, which is undesirable and can lead to additional heating or photo-bleaching problems.

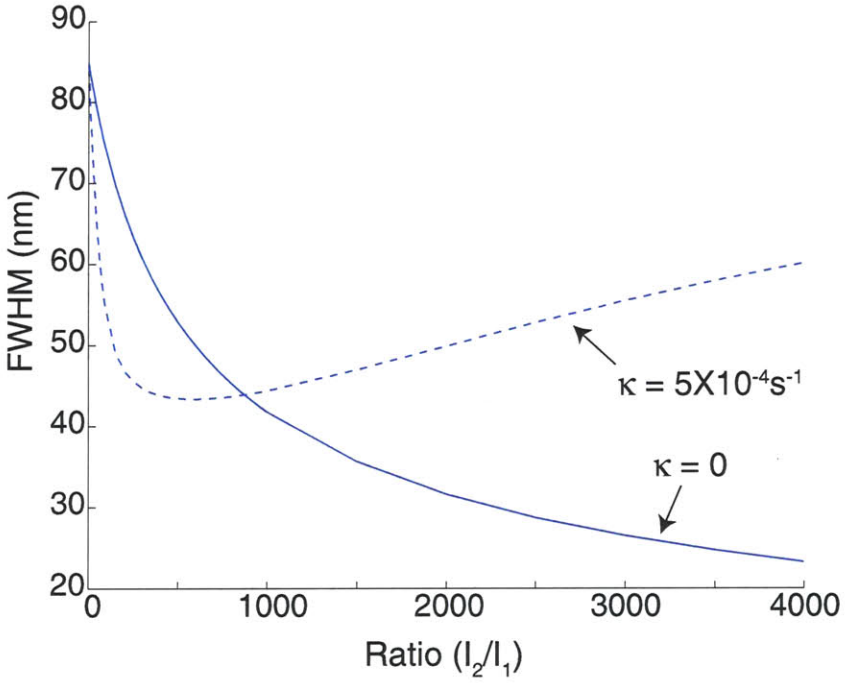

Figure 3-3: FWHM of the intensity distribution at λ_1 directly beneath the photochromic layer as a function of the ratio of the peak intensities at the two wavelengths. When the photochromic molecules are thermally stable [the thermal rate constant (κ) = 0], the size of the writing beam decreases monotonically, far below the wavelength. However, when a thermal instability is present ($\kappa = 5 \times 10^{-4} s^{-1}$), the smallest beam size is limited, as shown by the dashed line.

Fig. 3-4 shows the chemical formula of the azobenzene side-chains, a schematic of trans-cis isomer transitions, the corresponding transition parameters, and absorbance

measurements (at 405 nm) as a function of time for a 200 nm thick azobenzene AML. The subscripts A and B refer to the trans and cis isomers, respectively, while the subscripts 1 and 2 refer to λ_1 and λ_2 respectively. The symbols ϵ, ϕ, and κ refer to the corresponding molar absorption coefficient, the quantum efficiency and the thermal rate constant, respectively. As the azobenzene AML is thermally unstable and spontaneously converts from its cis-state (transparent state) to trans-state (opaque state,) the thermal rate constant κ_{BA} is measured to be nonzero, while κ_{AB} is negligible.

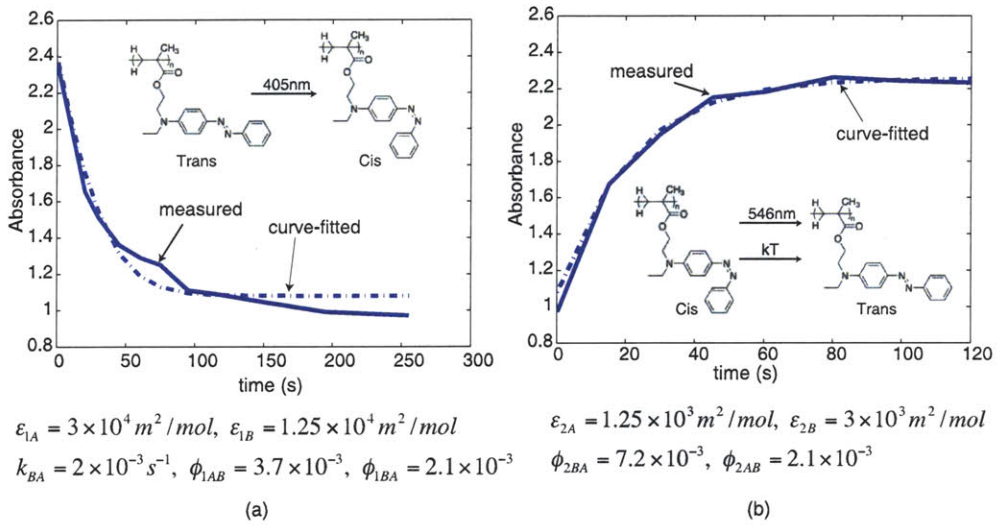

Figure 3-4: Absorbance (at 405 nm) measured as a function of time, when a 200 nm thick photochromic layer was illuminated by (a) 1.2 mW/cm^2 light at 405 nm, and subsequently by (b) 20 mW/cm^2 light at 546 nm. The experimental data (solid) are overlaid with the simulation (dashed), which enables the extraction of the photochromic parameters.

In a separate set of experiments, we showed that utilizing the linewidth compression and reversibility with the AML, features with spatial frequencies doubled from those present in incident intensity distributions can be patterned via a double-exposure, but single development process [97]. A schematic of the double-exposure technique is shown in Fig. 3-5. The same optical set up as in Fig. 3-1 was used. After the first exposure step at one fixed substrate angle with respect to the incident λ_2=532 nm beam, the substrate angle is changed slightly before the second exposure to expose with a slightly different grating period at λ_2=532 nm. There two exposures result in beat patterns in the photoresist, and spatial frequency doubling can be ob-

served at certain regions where one exposure happens to lie in between two lines from the other exposure, as shown in Fig. 3-6.

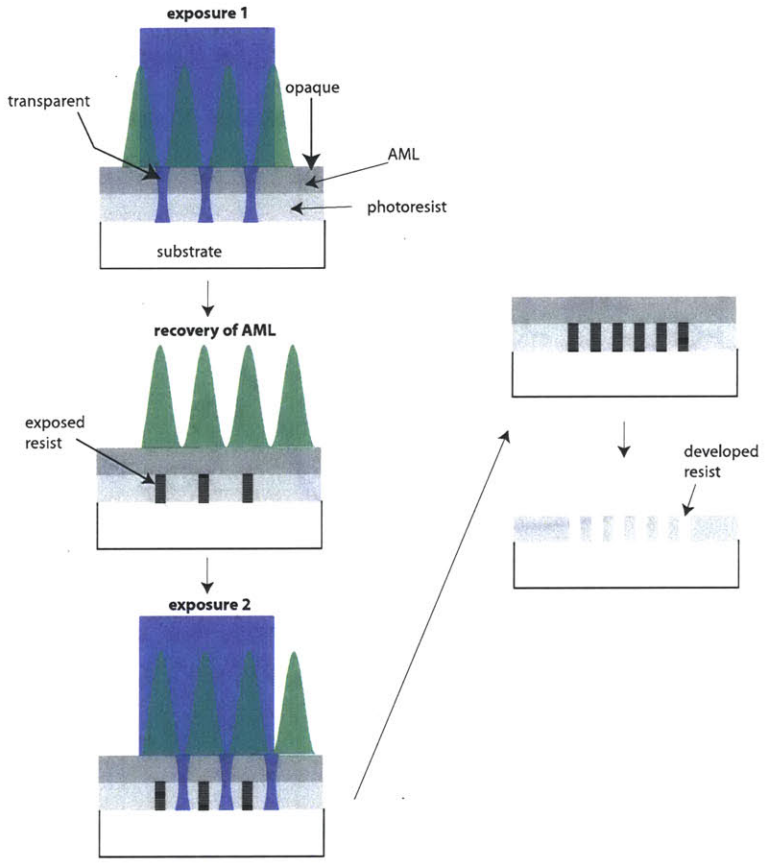

Figure 3-5: Schematic illustrating the double-exposure technique. The first exposure is conducted with a period, P_1 of the standing wave at λ_2. The period is changed to P_2 for the second exposure. A long exposure in between the two exposures with only λ_2 ensures that the AML recovers to its original opaque state.

These first demonstrations proved the feasibility of using the dynamically generated absorbance contrast in a thin film as the mask for lithography. With the reversibility of this absorbance modulation and proper engineering of the modulating far-field optics, high-resolution dense features can potentially be patterned. In the following sections of the chapter, I will show how we further approach this goal by improving the photochromic materials in the AML.

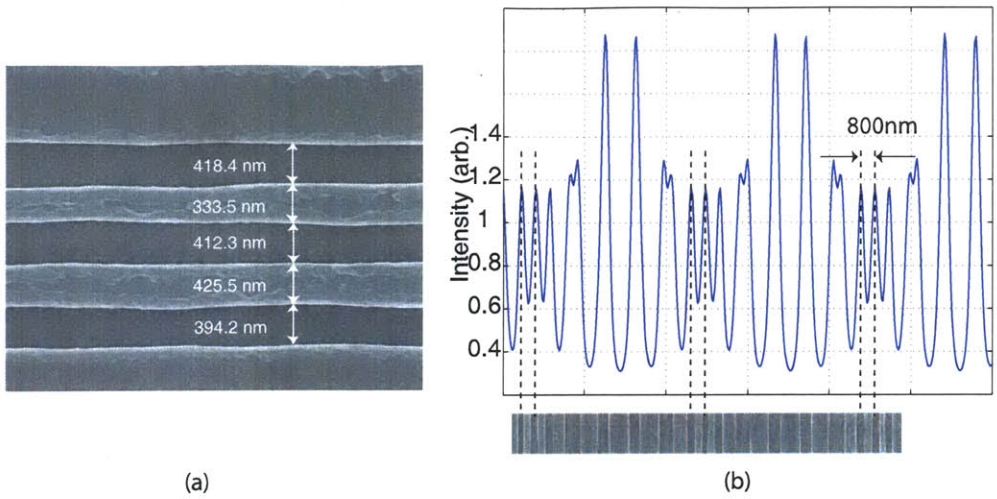

Figure 3-6: Spatial-frequency multiplication by double exposure. The first exposure was conducted at a period of the standing wave at λ_2 of 1.733 μm, and the period of the second exposure was 2.08 μm. (a) Scanning-electron micrograph of 3 lines. (b) The aerial image of the double-exposure pattern simulated by adding the two individual exposure patterns. Each exposure pattern was simulated using the model described in Chapter 2[92]. Note the agreement of the location of the fringes with the scanning-electron micrograph of the exposed pattern.

3.2 Previously Explored AMLs: Azobenzene Polymer, Fulgides, and Diarylethene

As described earlier, thermal instability of the AML limits how tight by light can be confined through absorbance modulation. In addition to the azobenzene AML described in the previous section, we also looked into two other candidates: the fulgide family and the diarylethene family.

In contrast to the azobenzene AML, which transitions between two geometric isomer states upon photon illuminations, the fulgide and diarylethene AML undergo photoinduced electrocyclic rearrangements and transform from a colorless (UV-absorbing) triene system to a highly colored cyclohexadiene photoproduct and vice versa. Because covalent bonds are either formed or broken during the photoisomerization process, conversion between the open-ring and closed-ring isomers is primarily photoinitiated, and the thermal contribution to this isomerization is negligible. Furthermore, the ring-closing transitions of these AMLs usually require photons at short

wavelengths in the UV range, while ring-opening can be photoinduced at similar or longer wavelengths than the λ_2 wavelength of azobenzene AML. This property can ease the search for proper photoresists because it further separates λ_1 and λ_2, and, therefore, makes it a lot easier to formulate photoresists that are exposed only by λ_1, but not by λ_2. On the flip side, UV illuminations are more energetic and can easily photo-bleache the AML by breaking bonds or inducing reactions that are not recoverable. Therefore, careful design of the molecule and proper protection from ambient is required for such AMLs.

The molecular structure and transition wavelengths of one member of the fulgide family are shown below. The structure and UV-Vis spectrum of the furylfulgide [98] is shown in Fig. 3-7 and Fig. 3-8. The furylfulgide molecule exhibits high absolute absorbance contrast, but revealed susceptibility to photodegradation that significantly reduced the concentration of this photochrome in the AML with prolonged irradiation. Cursory analysis of some fulgides reported in the chemical literature confirmed that many fulgides display a lack of fatigue resistance because of photooxidation of either their triene or heterocyclic moieties [99]. Although encapsulation slightly improved reversibility of the film, the protection did not last very long.

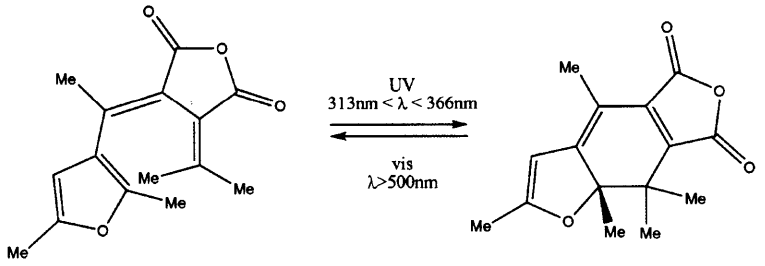

Figure 3-7: The molecular structure and transition wavelengths of the furylfulgide.

The chemical structure and UV-Vis spectrum of one diarylethene is shown in Fig. 3-9. Although Fig. 3-9 shows phototransitions at λ_1 and λ_2, lithographic patterns were not observed in photoresists when used as an AML. This may be due to the low absolute absorbance at λ_2, resulting in limited absorbance contrast. It is worthwhile to note that absorbance contrast is not only determined by the contrast provided by each single molecule, but also the volume concentration of the molecule. Unlike

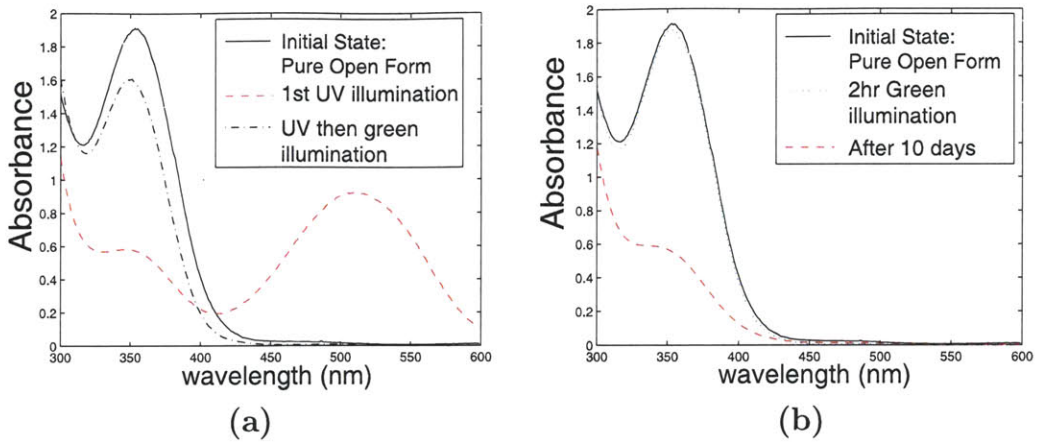

Figure 3-8: The UV-Vis spectrum of the furylfulgide. All fulgide samples were in PMMA and spun to 200 nm thickness. (a) The UV-Vis spectrum for the photostationary states after λ_1 illumination and the spectrum after λ_2 illumination with a 30nm PVA encapsulation layer spin-coated on the fulgide film. The reversibility of the UV-Vis curve is the result of the PVA protecting the fulgide molecules, increasing the photo-switching reversibility and lifetime of the underlying molecules. (b) The spectrum after long λ_2 illumination almost did not affect the spectrum. The dramatic change in absorbance after 10 days implies that fulgide is chemically unstable in the form of a film and has a short sample lifetime as a result.

the azobenzene AML, in which single molecule photochromes were polymerized and spun on in chloroform solvent, fulgide and diarylethene molecules were dissolved in a PMMA matrix to from films, so their concentrations are subject to solubility limits in PMMA.

3.3 The Compound 1

A comparatively photostable class of thiophenesubstituted fluorinated cyclopentenes was explored as potential photochromes for absorbance modulation. Specifically, the structure of 1,2-bis(5,5'-dimethyl-2,2'-bithiophen-yl) perfluorocyclopent-1-ene (compound 1) is shown in Fig. 3-10(a) and the absorbance spectra is shown in Fig. 3-10(b). Compound 1 displays an absorption band centered at 313 nm in the open state (opaque state to UV) and one centered at 582 nm in the closed state (transparent state to UV). These spectral features allowed the use of a 325-nm helium-cadmium (HeCd) laser and a 633-nm helium-neon (HeNe) laser as the patterning and the confining illu-

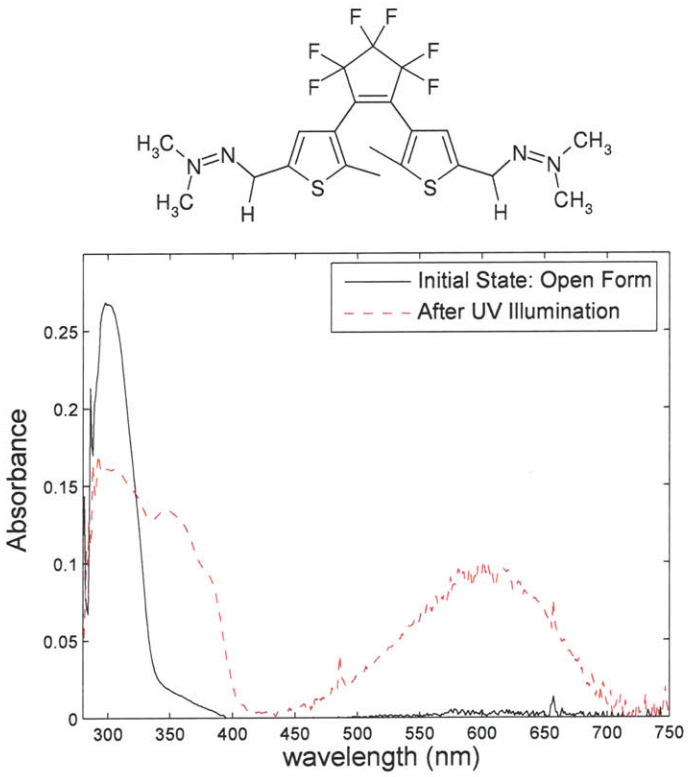

Figure 3-9: The molecular structure and UV-Vis spectrum of the diarylethene molecules used for AML. The diarylethene molecules were dissolved in PMMA and spun to 500 nm to increase the signal to noise ratio for the UV-Vis measurement. It can be seen that the diarylethene is clearly not very absorbing, resulting in an insufficient image contrast for the AML. λ_1=300 nm illumination was applied for the transition from open form to closed form, and λ_2=633 nm illumination was applied for the reverse transition.

minations, respectively. High intensities could be applied at the nodal wavelength λ_2 because 633-nm light has no effect on most photoresists. Photochromic parameters, such as absorption coefficients and photoreaction quantum yields, were measured for compound 1 at room temperature in hexane solution and shown in Table 3.1. Compound 1 shows a contrast at λ_1 between the open and close states of about 3, slightly better than that of the azobenzene AML. A negligible value of κ confirms that Compound 1 is indeed thermally stable. However, compound 1 also has some drawbacks as photochromes for absorbance modulation. The extinction coefficient at λ_1 for the opaque state is one order smaller than that of the azobenzene AML and, more importantly, the quantum yields of the forward and backward photoreactions are highly unbalanced. The quantum yield of the ring-closing reaction is 300 times higher than that of the ring-opening reaction. This means that a much more extreme λ_2 to λ_1 intensity ratio would be required to achieve the same level of confinement to compensate for that yield difference.

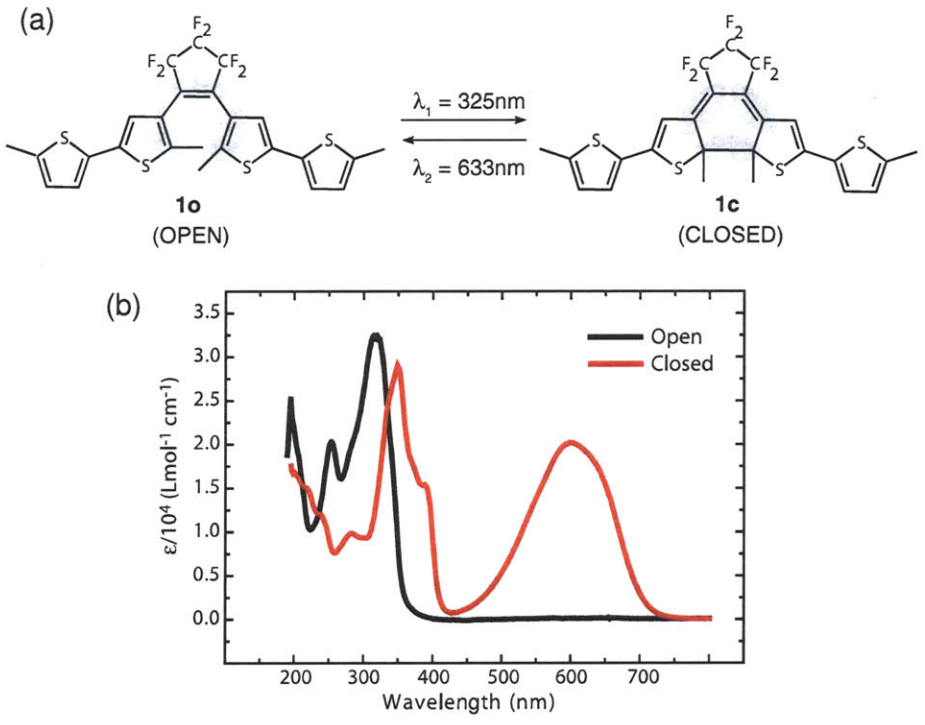

Figure 3-10: (a) Structures of the open- and closed-ring isomers of compound 1. (b) Absorbance spectra of compound 1 in the open and closed forms in hexane. ϵ is the decadic molar absorptivity.

Table 3.1: Photochromic parameters for compound 1 and azobenzene AML at room temperature.

Compound 1	Azobenzene AML
$\epsilon_{1open} = 3.11 \times 10^3$ m^2/mol	$\epsilon_{1A} = 3 \times 10^4$ m^2/mol
$\epsilon_{1closed} = 1.05 \times 10^3$ m^2/mol	$\epsilon_{1B} = 1.25 \times 10^4$ m^2/mol
$\epsilon_{2open} = 15.8$ m^2/mol	$\epsilon_{2A} = 1.25 \times 10^3$ m^2/mol
$\epsilon_{2closed} = 2.00 \times 10^3$ m^2/mol	$\epsilon_{2B} = 3 \times 10^3$ m^2/mol
$\Phi_{open->closed} = 0.24$	$\Phi_{1A->B} = 3.7 \times 10^{-3}$, $\Phi_{2A->B} = 2.1 \times 10^{-3}$
$\Phi_{closed->open} = 0.24$	$\Phi_{1B->A} = 2.1 \times 10^{-3}$, $\Phi_{2B->A} = 7.2 \times 10^{-3}$
$\kappa < 3 \times 10^{-6}$ s^{-1}	$\kappa_{B->A} = 2 \times 10^{-3}$ s^{-1}, $\kappa_{A->B} \sim 0$
$[\text{open}_0] = 2.99 \times 10^3$ mol/m^2	$[\text{A}_0] = 400.024$ mol/m^2

The synthesis and photochromic characterization were contributed by Trisha L. Andrew, PhD student from Prof. Timothy M. Swager's group. The photochromic characterization process is described in more detail below.

The molar absorption coefficients of the open and closed forms of compound 1 at 325 nm (λ_1) and 633 nm (λ_2) (ϵ_{1open}, $\epsilon_{1closed}$, ϵ_{2open}, $\epsilon_{2closed}$) were calculated using the integrated form of the Beer-Lambert Law ($\alpha = \epsilon lc$, where α is absorbance, l is the optical path length, and c is the concentration.) The law states that there is a logarithmic dependence between the transmission (or transmissivity), T, of light through a substance and the absorbance.

$$T = \frac{I}{I_0} = 10^{-\alpha} \tag{3.1}$$

The ultraviolet-visible absorbance values were measured with an Agilent 8453 diode array spectrophotometer and used in the integrated Beer-Lambert equation to calculate the experimental absorption coefficients. The samples consist of a known quantity of pure compound 1 (open) and compound 1 (closed) dissolved in hexanes in the dark. Approximately 3 mL of these two solutions were placed in quartz cuvettes with an optical path length of 1 cm when the absorbance spectrum recorded.

The quantum yield for ring-closing was defined as the moles compound 1 (closed) formed per mole photon absorbed by the sample. Ring-closing quantum yields were obtained by irradiating hexane solutions (of known concentrations) of compound 1 (open) at 313 or 325 nm and monitoring the change in absorbance at 633 nm at 5s intervals. The quantum yield for ring-opening was defined in a similar fashion, and obtained by irradiating solutions of compound 1 (closed) with 633 nm light and monitoring the change in absorbance at 313 nm. The absorption profile of the ring-opening photoreaction was recorded every 50 minutes. A 500 W Xenon arc lamp with monochromators at corresponding wavelengths was used as the light source. The irradiation density was measured using a power meter and converted to moles of photons per second using Einstein's equation ($E_{mol} = N_A h\nu$, where N_A is Avogadro's number). The moles of photons absorbed by compound 1 was calculated from the irradiation duration and corrected with the absorbance immediately prior to irradiation.

The initial concentration of the open form as an AML was calculated using the integrated form of the Beer-Lambert law with the absorbance measured from a 410 nm-thick film. In order to spin-cast the photochromic layer, we used a 30 mg/ml solution of poly(methyl methacrylate) (PMMA) in anisole doped with 92 weight percent compound 1 (with respect to PMMA) [100]. The film was spun-cast under yellow light in a class 10 clean-room and kept in the dark before absorbance measurements. The dispersion curve and thickness of the film were measured using a spectroscopic ellipsometer and the absorbance of this film was calculated using the equation, $\alpha = 4\pi k/\lambda$, where k is the imaginary part of the refractive index.

In order to check the thermal stability of compound 1, solutions of open and closed form compound 1 in toluene were heated to 100°C in the dark for 60 min and their absorbance spectra were recorded every 10 minutes. Additionally, thin films of 92 wt% compound 1 in open and closed form in PMMA were heated to 120°C for 60 minutes and their absorption spectra recorded every 10 minutes. No changes in absorbance spectra were observed within the measurement error of our spectrophotometer (~0.01 optical density, O.D.). Considering the 60 mins of observation time, we conclude that

the thermal reaction rate constant, κ, is less than 3×10^{-6} s^{-1}. For the sake of simplicity, κ is set to 0 s^{-1} for all the simulations, unless stated otherwise.

3.4 Interference Lithography with Compound 1

The exposure system was a modified Lloyd'smirror interferometer, as shown in Fig. 3-11. This illumination wavelengths were $\lambda_1 = 325$nm for ring-closing (opaque to transparent transformation) and photoresist patterning, and $\lambda_2 = 633$ nm for ring-opening. The angles of incidence of the two wavelengths were adjusted so that the resulting standing-wave period at λ_2 is two times the period at λ_1. As illustrated in Fig. 3-11, the nodes of the λ_2 standing wave approximately coincide with every other peak of the λ_1 standing wave. Photokinetic simulation using the extracted photochromic parameters reveal that the transmitted light at λ_1 is substantially narrower than the diffraction limit.

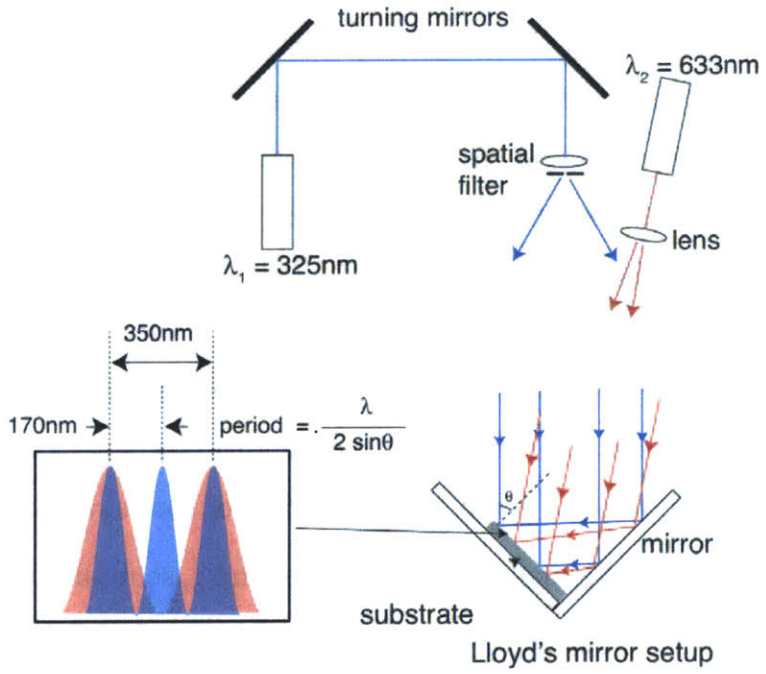

Figure 3-11: Schematic of the dual-wavelength Lloyd's-mirror interferometer. The inset shows the interlaced pair of standing waves incident on the substrate. The angles of incidence of the two wavelengths are adjusted so as to form standing waves of period 350 nm at λ_2 and 170nm at λ_1.

3.4.1 The Lloyd's Mirror Setup

A schematic of the dual-wavelength Lloyd's-mirror interferometer is shown in Fig. 3-11. A Lloyd's-mirror setup consists of a mirror at right angles to a vacuum chuck that holds the sample. When half of the incident beam is reflected by the mirror onto the sample surface, the reflected beam interferes with the remaining portion of the incident beam (that is directly incident on the sample) and produces a standing wave. The angles of incidence of the two wavelengths were adjusted so that the resulting standing waves on the sample had periods of 350 nm at λ_2 = 633 nm and 170 nm at λ_1 = 325 nm. Light from a helium-cadmium laser (λ_1 = 325nm) is spatial-filtered and expanded over a distance of about 1 meter. This generates an approximately uniform plane wave at the Lloyd's mirror. Light from a helium-neon laser (λ_2 = 633 nm) is introduced at an angle as shown. Because the period of the λ_2 gratings is not exactly two times the grating period at λ_1, relative phase between the two sets of gratings beats. In some regions, the peaks of the λ_1 standing wave overlap with nodes of the λ_2 standing wave, creating the ideal illumination for absorbance modulation in one dimension. Polarization of both the beams was ensured to be pointing into the plane of the figure to maximize the contrast of the standing waves. The intensity at λ_1 as measured at the Lloyd's mirror was 0.365 W/m^2. The λ_2 beam was expanded perpendicular to the axis of the Lloyd's mirror via a cylindrical lens placed at a distance of 61 cm from the Lloyd's mirror. The peak intensity of the illumination at the Lloyd's mirror was estimated to be about 1×10^3 W/m^2.

3.4.2 Sample Preparation and Development Process

The recording sample consists of a stack as shown in Fig. 3-12 on a silicon substrate. A silicon wafer was spin-coated with 200nm of BarLi anti-reflection coating (ARC) and then with a photo-sensitized chemically amplified positive-tone KrF photoresist that is formulated in house. Two resist formulas were tested for exposures at the λ_1 = 325 nm wavelength: one comprised of a 200nm-thick layer of TDUR-P308 (Tokyo Ohka Kogyo Co. Ltd.) resist with an i-line (365nm) photoacid generator, CGI-725

(CIBA Specialty Chemicals, Inc.,) in the ratio of approximately 5:1 by weight, and the other comprised of a 73nm-thick layer of TDURP-P308 with another i-line photoacid generator, Irgacure PAG-103 (CIBA Specialty Chemicals, Inc.), in the ratio 20:1 by weight. Photoacid generators sensitize the KrF resist by generating acids that convert alkaline-insoluble polymers into alkaline-soluble polymers upon absorption of photons at specific wavelengths. Most photoacid generators are proprietary, but the structure and spectrum of the Irgacure PAG-103 series are shown below in Fig.3-13. As shown in the spectrum, PAG-103 provides limited sensitization at $\lambda_1 = 325$ nm and was first tried for a fomulatation that is only sensitive to 405 nm, but not to 532 nm light (for the azobenzene AML.) Therefore, the formulation with CGI-725 was chosen for most of the experiments with compound 1. Chemical amplification in photoresists were required in these experiments to record highly attenuated λ_1 illuminations, which were required to achieve high enough λ_2 to λ_1 ratio. The requirement of a high λ_2 to λ_1 ratio is due to the imbalance nature between ring-opening and ring-closing reactions of Compound 1. This ratio is ultimately limited by available λ_2 intensity and sensitivity of the photoresist at λ_1.

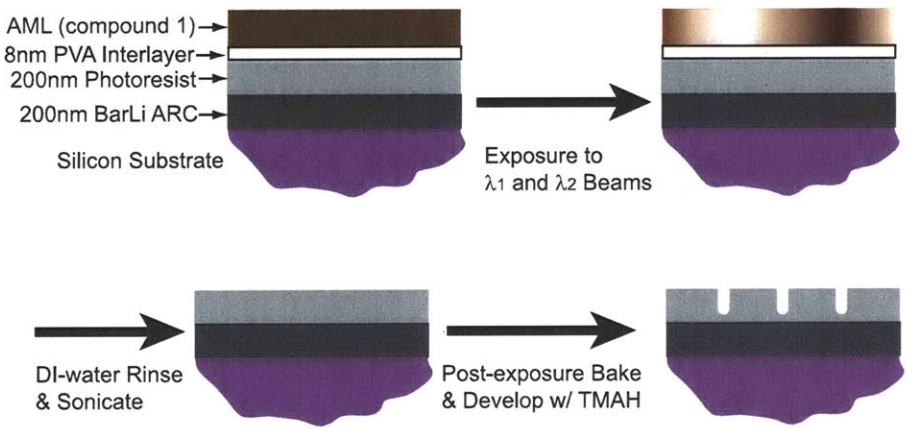

Figure 3-12: The recording stack and development process for Lloyd's mirror exposures with absorbance modulation.

Before spin-coating AML (compound 1) on top of the photoresist stack, the sample was spin-coated with a solution of poly-vinyl alcohol (PVA) in water (5.7 μg/ml) and air-dried for 5 minutes to form a film of thickness 8nm. Because the solvent

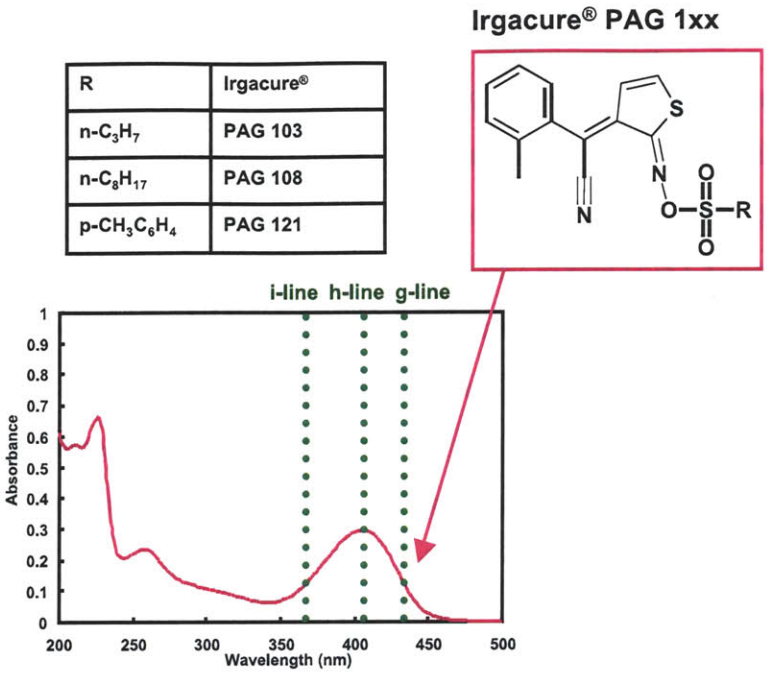

Figure 3-13: Spectrum of one of the photoacid generator (PAG-103) used for resist formulation.

for the PMMA matrix, anisole, distorts the development rate of the photoresist, a barrier layer was required in between the two layers. The barrier layer also prevents any interdiffusion between the two layers. Nevertheless, the high-spatial frequency contents of the nanoscale writing profiles evanescently decay upon exiting the bottom of the AML. As a result, it is important to keep the thickness of the PVA layer as small as possible to minimize the effect of line broadening. We found that 8nm of PVA was sufficient to protect the photoresist from the solvent of the AML. Finally, the sample was spin-coated with the PMMA solution doped with compound 1 and air-dried for 5 minutes to form a 410nm-thick AML.

After exposure, the samples were rinsed in de-ionized water in a sonicator for about 5-10 minutes, which removed the PVA layer as well as the AML. The photoresist was baked on a hotplate at 120oC for 90 s and developed in 0.26 N tetramethyl ammonium hydroxide (TMAH) for 60 s. The resulting patterns were inspected in a scanning electron microscope (SEM) with 2 nm sputter-coating of Au/Pd alloy.

3.4.3 Lithography Results

In some regions of the Lloyd's-mirror interferometer, the standing waves at λ_2 and λ_1 overlap such that each node at λ_2 approximately coincides with each alternate peak at λ_1. Those peaks at λ_1 that coincide with the peaks at λ_2 are suppressed beyond the AML. Figure 3-14(a) shows the SEM image of a suppressed grating pattern where the intensity ratio at λ_2 over λ_1 is not high enough. In these regions, the "suppressed λ_1 peaks" are weakly recorded in the underlying photoresist. This effect is visualized more clearly in Fig. 3-14(b), where the SEM image of a cross-section through the lines recorded in photoresist reveal that the "suppressed peaks" are indeed shallower than the others.

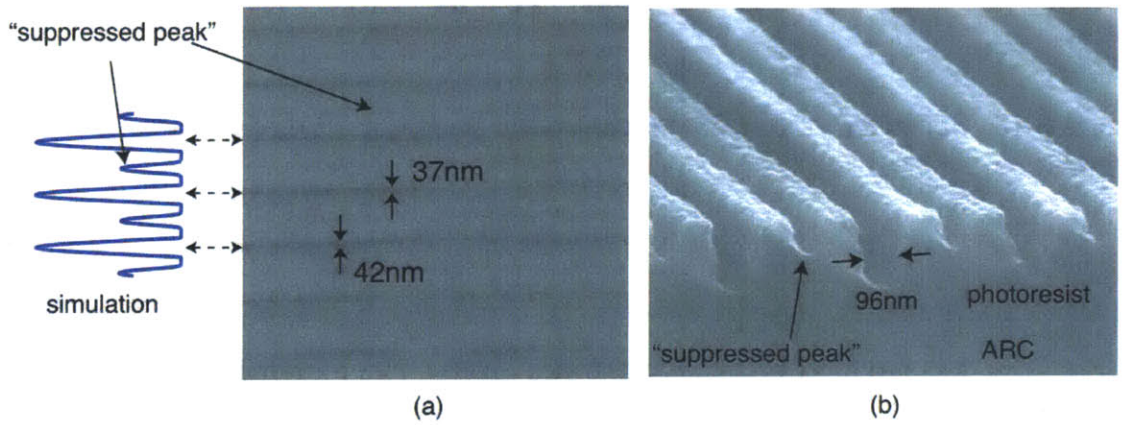

Figure 3-14: (a) Top-down SEM image of a region where the standing waves at λ_2 and λ_1 overlap such that each node at λ_2 approximately coincides with each alternate peak at λ_1. Due to the relatively low intensity ratio at λ_2 over λ_1 (compared to the ratio required for tight confinement with compound 1), "suppressed peaks" at λ_1 are weakly recorded in the photoresist. (b) SEM image of a cross-section through exposed lines in photoresist from the sample. The "suppressed peaks" are correspondingly shallower than the other peaks. The period of the main peaks in both micrographs is equal to 350 nm, which is the period of the λ_2 standing wave.

Figure 3-15(b) shows the SEM image of a portion of the pattern on the same sample as in Fig. 3-14(b), but at a different location with more intense λ_2 illumination. The SEM image in Fig. 3-15(b) shows that the average width of the lines recorded in the photoresist was 36 nm, close to one tenth of λ_1. Furthermore, the narrow lines

were spaced by 350 nm, which corresponds to the period of the λ_2 standing wave. We separately confirmed that the photoresist is not sensitive to the λ_2 photons, so that we confirmed neither the suppressed grating lines in Fig. 3-14 nor the highly confined grating lines in Fig. 3-15 were due to λ_2 exposure.

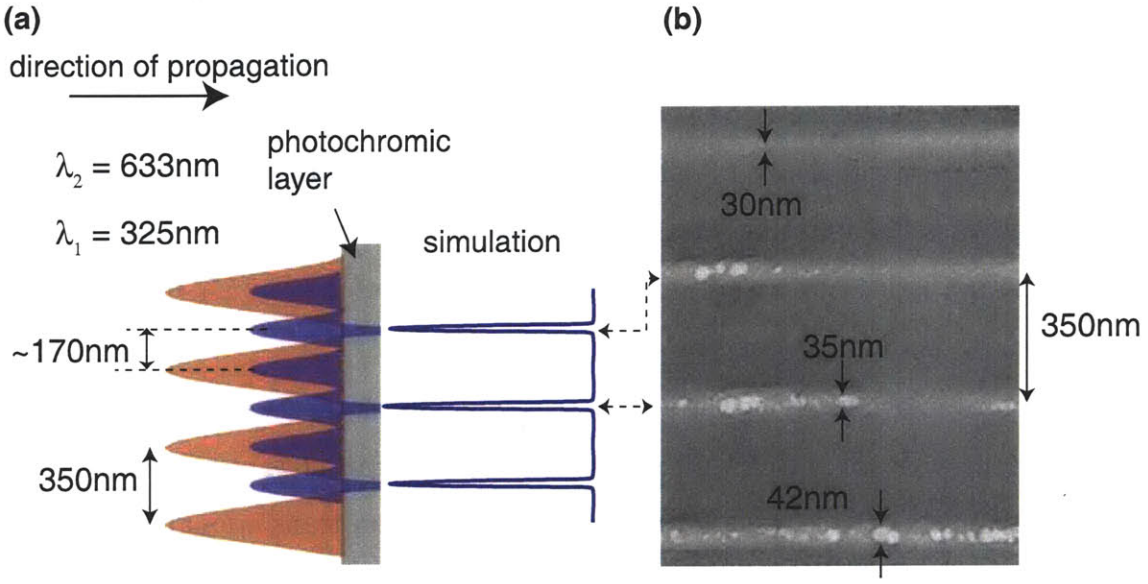

Figure 3-15: Deep subwavelength patterning using absorbance modulation. (a) The photochromic layer is illuminated by two overlapping standing waves with periods of 350 nm ($\lambda_2 = 633$ nm) and 170 nm ($\lambda_1 = 325$ nm), respectively. Simulating the transmitted light at λ_1 supported narrow lines where the peaks of the λ_1 standing wave coincided with the nodes of the λ_2 standing wave. (b) An SEM image of lines exposed in photoresist. Although the photoresist is underexposed, the lines represent a recording of the aerial image that is consistent with simulation.

In our current setup, in order to achieve high intensity in the λ_2 peaks, it was necessary to remove spatial filtering for the λ_2 illumination. High-frequency noise therefore was present in the λ_2 standing wave, causing line edge roughness as well as the line-width variation in the photoresist patterns. Nevertheless, these results clearly demonstrate the feasibility of deep subwavelength localization of light by using absorbance modulation. Furthermore, these results also demonstrate the feasibility of patterning periodic lines far smaller than their spatial period. Because the absorbance of the AML is reversible, interspersed multiple exposures could pattern lines spaced apart by a distance far smaller than the far-field diffraction limit of the optical system

[97].

3.5 Summary and Future Work

In this chapter, Absorbance Modulation experiments with Lloyd's mirror (LM) interference lithography were reviewed and presented. With a simple experimental setup, we studied how photokinetic properties of the AML can greatly affect AML output profiles and investigated several different classes of photochromic materials as the AML. Highly confined grating lines as narrow as 40 nm were recorded in photoresists with a 325 nm light source and an AML that consists of 1,2-bis(5,5'-dimethyl-2,2'-bithiophen-yl) perfluorocyclopent-1-ene (compound 1.)

In the future, by modifying the system to conduct experiments with double or even more exposures, dense patterns of highly confined lines can be achieved. One way to do so is by placing the Lloyd's mirror on a high precision translational stage. After one exposure, the LM setup can be translated by half the grating period (85 nm in the above case) in the direction parallel to the substrate for the second exposure. Another way is by replacing the Lloyd's mirror setup with a Mach-Zehnder interferometer. In addition to usually giving better beam uniformity, grating patterns created with the Mach-Zehnder interferometer can be accurately shifted by controlling the phase shift of one arm relative to the other without physically moving the sample.

The absorbance modulation layer can also be used as a reversible contrast enhancement layer (RCEL) for conventional lithography processes with a single wavelength. The concept of contrast enhancement for lithography was first proposed in 1990 [94] targeting applications for integrated circuit production. The contrast enhancement layer (CEL) consists of a photo-bleachable photochromic material which becomes transparent upon illumination at the exposure wavelength. Because high intensity profiles bleaches through the CEL first (as demonstrated in Chapter 2), higher intensities experience lower attenuation while lower intensities experience higher attenuation. This effectively increases contrast of the illumination profile. When this photo-bleaching is reversible, the CEL can be recovered to its initial state for subse-

quent exposures without removing the film. This technique is referred to as double exposure and has been widely studied as an option for frequency doubling [101]. Since RCEL materials are similar to AML in many aspects, studies on feasibility of double patterning materials can also be relevant for absorbance modulation [102].

Chapter 4

Generation of Optical Nodes for Absorbance Modulation

In order to extend Absorbance Modulation from 1D-confinement to 2D-confinement, *i.e.* from patterning grating lines to patterning round-shaped spots, an illumination scheme, in which a focused round spots at one wavelength, λ_1, is aligned with the central node of a ring-shaped spot at another wavelength, λ_2, is required. The quality of the node and alignment between the spot and the node are critical parts of techniques that break the farfield diffraction barrier, including stimulated emission-depletion (STED) fluorescent microscopy [103] and absorbance modulation [96]. Ring-shaped spots with central nodes are typically generated using a wavefront modulator [104] or by focusing Laguerre- Gaussian modes [105] and higher-order Bessel beams from computer-generated holograms [106, 107]. In those cases, an additional step is required to align a focal spot at λ_1 to the ring-shaped spot at λ_2. This alignment step might be feasible for a single pair of focal rings and focal spots, but would be difficult for arrays of rings and spots, which are of interest for massively parallel high throughput microscopy and lithography [108]. Single elements, such as spiral phase plates [109] or spiral zone plates [110], could be used to form ring-shaped spots at λ_2, but the fabrication of such elements is quite complicated, the resulting phase profile is very sensitive to fabrication errors, and the focusing at λ_1 would be suboptimal.

In this chapter, I will describe the design and fabrication process of a diffractive

optical lens, called dichromat, that simultaneously generates a focal spot at λ_1=405 nm, coincident with a ring-shaped spot at λ_2=532 nm. The choice of wavelengths is motivated by the specific Absorbance Modulation Layer (AML) we used. Optical characterization of dichromats is shown at the end of the chapter, while the actual application of dichromats in absorbance modulation to demonstration of point-spread-function (PSF) compression will be shown in the next chapter.

4.1 Diffractive-Optical-Element Design

Diffractive-optical elements (DOE) are thin planar elements that can sculpt the propagation of light by imposing particular phase and intensity patterns on the incident light. As the name implies, the function of DOEs relies primarily on diffraction or interference rather than refraction. Popular examples of DOE's include Fresnel zone plates and holograms. Compared to conventional refractive optics, the fact that DOE's can be made on thin sheets greatly facilitates utilization and integration of DOE's with planar semiconductor-fabrication technology. This implies not only that the DOE's can be miniaturized with existing high resolution, high yield processes, but also that a large array of such DOE's can be fabricated in parallel. This is true especially for binary-phase DOE's, which contain only two levels of phase, and, therefore, are significantly easier to fabricate. In addition to the ease of fabrication and potential for parallelization, DOE's generally enable more complex designs of the output illuminations, however, at the expense of output efficiency. Unlike conventional refractive optics, DOE's generally suffer lower efficiencies due to absorption or unwanted diffractive orders. While absorption can be avoided by simply using pure-phase DOE's, unwanted diffractive orders are much more difficult to eliminate and often require specific cancellation through interference or through multi-level phase design. Since fabrication of multiple phase levels is difficult and may introduce more loss due to phase error, in this work, only binary-phase DOE's are considered.

Another common drawback of DOEs is that such optics inherently exhibit chromatic aberrations. However, there have been several approaches to design multiple-

wavelength DOE's. A heterogeneous design, based on materials with differing refractive indices and dispersion to compensate for chromatic aberration, was proposed [111]. By using phase shifts that are integer multiples of 2π, harmonic diffractive lenses can be designed for specific discrete wavelengths [112]. By blazing diffractive structures for higher-order diffraction, the design of achromatic singlets is enabled [113]. Furthermore, in this particular case of dichromat design, chromatic aberrations can be used to our advantage to optimize for different illumination patterns at the plane of interest.

In this section, we describe a technique that utilizes nonlinear optimization to design lenses that can focus several wavelengths of light into focal spots of distinct geometries. In particular, we design a dichromat, a lens that focuses one wavelength, λ_1, to a bright spot and a second wavelength, λ_2, to an overlapping ring-shaped spot. The latter, with a node in its center, is generated in the absence of any phase singularities. As opposed to nodes generated via phase singularities, we demonstrate that the focal null of the dichromat can be made tolerant to fabrication errors. To further ensure ease of fabrication, these lenses are composed of concentric circular zones. The ultimate objective is to fabricate an array of such dichromats to enable parallel superresolution patterning and imaging [103, 96].

4.2 Defining the Problem: Generating Optical Nodes with Circularly Symmetric Elements

The design of dichromats is a process that optimizes certain figure of merits, usually in the output profiles, under given constraints. In this design problem, the figure of merit to maximize is the focusing efficiency and quality of λ_1 to a round-shaped spot and λ_2 to a ring-shaped spot at a given focal plane. The constraints include assuming a binary-phase circular-symmetric element, given dispersion characteristic of the phase object, a given focal distance, and the ease of fabrication (constraint for minimum feature size). Note that neither the figures of merit nor the constraints

are strictly defined or required. The efficiency and quality of the output illumination can be defined and weighted differently in the optimization process. In this work, various definitions for the optimization targets will be discussed while keeping most constraints the same. The chosen optimization technique was originally proposed by di Francia [114], where the optic is composed of concentric circular zones whose radii are the design variables. In some designs, the phase shift between adjacent zones was treated as an additional degree of freedom. Although many of the constraints can potentially be eliminated or loosened to enable even larger search space for optimal solution, parameterizing such element would require more degree of freedom and additional efforts to ensure fabrication feasibility. For example, eliminating the requirement of a circular-symmetric element would widen the solution space to include solutions similar to a spiral zone plate, which in principle can give true nodes on the optical axis, but would require a different set of parameters from radii of concentric rings. Finding a set of parameters that can include all types of designs is very difficult, so the design was limited to binary-phase circular-symmetric dichromats throughout this work.

4.2.1 Design Variables

The dichromats can be described by the following circular-symmetric transmission function:

$$T(\rho) = \begin{cases} e^{i\psi}, & r_{2m} < \rho \leq r_{2m+1},\ m \in 0, 1, 2, ... \\ 0, & \rho > r_M \\ 1, & \text{elsewhere} \end{cases} \tag{4.1}$$

where ρ is the radial coordinate, r_m is the radius of the mth zone, and M is the total number of zones. The relative phase shift between neighboring zones, ψ, can be related to the zone height, h, via

$$\psi = 2\pi \frac{h}{\lambda}[Re(n(\lambda)) - 1], \tag{4.2}$$

where $Re((n))$ is the real part of the refractive index of the lens material. The design parameters are the radii of the zones, r_1 ,r_2, $\cdots$,r_{M-1}, and the height of the zones, h. A cross-sectional schematic of the DOE and the design parameters are shown in Figure 4-1.

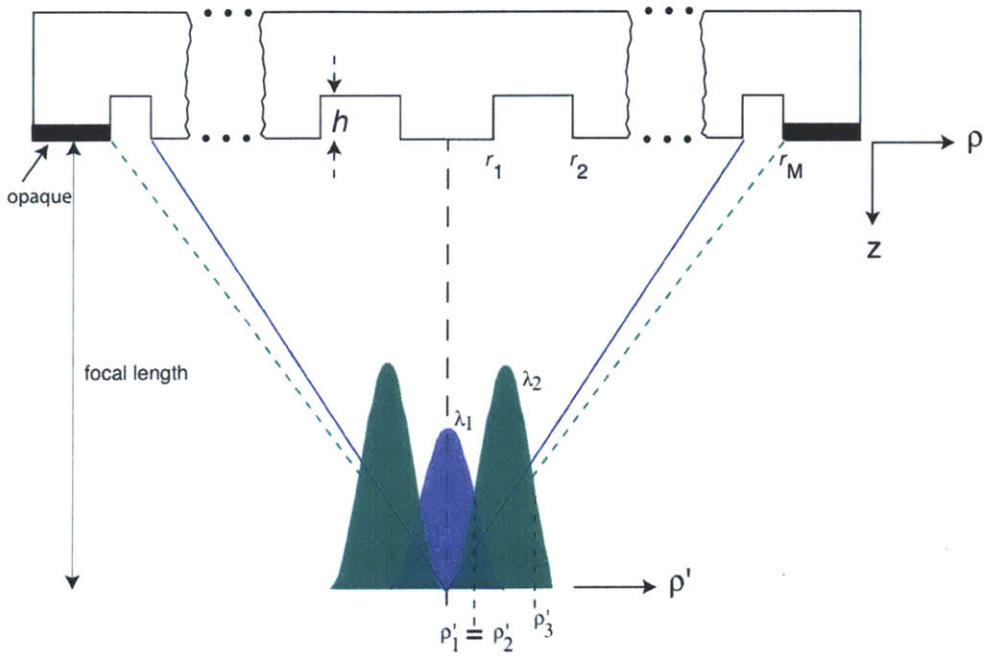

Figure 4-1: A cross-sectional schematic of the DOE showing the design parameters.

The dichromats were designed with a focal length of $100\lambda_1$=40um and numerical apertures of NA=0.56 (40 zones), NA=0.7 (80 zones), and NA=0.83 (160 zones). The design wavelengths were λ_1=405 nm and λ_2=532 nm. r_M was fixed at the radius that would give the corresponding numerical aperture (NA) for the given focal length. Fresnel-Kirchoff formulation of the scalar diffraction problem was used to model the propagation of light from the optic to the plane of observation [66]. A normally incident uniform plane wave is assumed.

The intensity at the focal plane is then given by

$$I(\rho', z, \lambda, \{\overrightarrow{r}, h\}) = |\frac{1}{i\lambda}\int_0^\infty \rho d\rho \int_0^{2\pi} T(\rho, \{\overrightarrow{r}, h\}) \frac{\exp i\frac{2\pi}{\lambda}\sqrt{\rho^2 + \rho'^2 - 2\rho\rho'\cos\phi + z^2}}{\sqrt{\rho^2 + \rho'^2 - 2\rho\rho'\cos\phi + z^2}} \tag{4.3}$$

$$(1 + \frac{z}{\sqrt{\rho^2 + \rho'^2 - 2\rho\rho'\cos\phi + z^2}})d\phi^2,$$

where ρ and ϕ are the cylindrical coordinates in the plane of the diffractive lens, z is the propagation distance along the optical axis, and rho' is the radial coordinate in the observation plane. The scalar-diffraction theory was used for its high computation speed. However, the proposed method is equally applicable with full-vector diffraction theories. A rigorous full-vector electromagnetic theory analysis was developed by Dr. Rejesh Menon to verify the focusing performance of some dichromat designs to confirm applicability of the scalar-diffraction theory.

4.2.2 Additional Design Constraints

In addition to fixed design inputs, some additional constraints on the dichromat radii are imposed to ensure fabrication feasibility:

$$r_p > r_q \forall M \geq p > q \geq 1, \{p, q\} \in I, \tag{4.4}$$

$$r_p - r_{p-1} > \Delta > 0 \forall M \geq p > 1, p \in I, \tag{4.5}$$

$$r_1 > 2\Delta \tag{4.6}$$

The first constraint ensures that the zones are retained in the correct order during optimization. The last two constraints ensure that the width of each zone is greater than the minimum feature size, delta, a value dictated by fabrication technology. In this work, delta was set to λ_1=405 nm for the 40-zone dichromat and to $0.5\lambda_1$=200nm for all other dichromats. The fabrication technology, as will be described in more detail later, is direct scanning-electron-beam lithography (SEBL) on the phase shifting material, *i.e.* PMMA or hydrogen silsesquioxane (HSQ). The resolution of SEBL has been demonstrated to be as high as 14nm pitch [ref], but in thin HSQ resists with

a high contrast development process. Because the height of the zones, h, in most dichromat designs lies around 400 nm, a delta of 200 nm or larger was found to provide reasonable process tolerance and adequate feature aspect ratios.

4.3 Optimization Method

One key step of the design process is the optimization algorithm. The goal of the optimization is first described in terms of a fitness function and then maximized to find optimal solutions. Since Eq. 4.3 is a highly nonlinear function of the spatial coordinates and the fitness function and the constraints add additional complexity, we chose genetic algorithm for this nonlinear optimization problem.

4.3.1 The Genetic Algorithm

The genetic algorithm is an iterative mathematical version of natural selection [115, 116]. In this work, we used the Genetic Algorithm (GA) toolbox in Matlab to implement the optimization process. The GA begins with an initial population, which consists of 25 or 50 individuals, each representing a distinct design of the diffractive lens. The number of individuals, referred to as the population size, is a variable option set before calling GA and remains constant over iterations. The GA then iteratively generates a series of new populations through mutation and crossover. Mutation means adding random perturbations to the randomly selected individuals in the current generation, while crossover means randomly selecting complimentary sections from pairs of individuals in the current generation and combining them. There are several options for mutation methods implemented in the Matlab GA toolbox. The mutation method used for this work was "Adaptive Feasible", which randomly generates directions that are adaptive with respect to the last successful or unsuccessful generation. The crossover fraction, *i.e.* the fraction of the population at the next generation that is created by the crossover function, is set to 0.4. The GA options are summarized in Table 4.1.

At the end of each iteration, every individual is ranked based on the value of a

Table 4.1: Options for the GA function in Matlab.

GA Options	Setting Example	Comment
PopulationSize	25 or 50	Population size
StallTimeLimit	Inf	If no improvement in this time, quits
Display	Iter	Displays each iteration
MutationFcn	@mutationadaptfeasible	Adaptive Feasible Mutation Function
InitialPopulation	Array of population	Size of array is number of parameters by population size
Generations	25-100	Stop after maximum number of iterations
CrossoverFraction	0.4	Fraction of next generation that is created by crossover
Vectorize	On	Computes entire population at the same time to increase speed

fitness function. The individuals for the next iteration are generated by retaining those in the current generation with the lowest fitness values up to the size of the original population. During each iteration, individuals are also forced to satisfy all the linear constraints and bounds. The constraints can be imposed by defining two input variables A and b of the "ga" function in Matlab to force all individuals in the population to satisfy the inequality contraints: $A * x \leq b$. In this problem, a typical A and b matrix would look like the following,

$$\begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} * \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_M \\ h \end{bmatrix} \leq \begin{bmatrix} -\Delta \\ -\Delta \\ -\Delta \\ \vdots \\ R_n \\ 1 \end{bmatrix}, \tag{4.7}$$

where R_n is the radius of the outer most zone (same as r_M), defined by the NA or number of zones of the dichromat. In some cases where h fixed and, therefore, not included as a parameter, the last line in the inequality constraints is removed. The lower bound constraint Eq. 4.6 is imposed by setting the LB parameter of the GA function, with which $LB \leq x$ is enforced. This iterative process is continued until a

stopping criterion is met (no improvement upon iteration) or when maximum number of iterations is reached.

4.3.2 Initial Population and Evolution

For optimization with GA, the initial population can greatly impact the speed of convergence and the quality of the final solution. Since GA relies on random mutation and crossover from the initial population, starting with a good guess could save a lot of iterations for generating the right children from evolution. One intuitive starting guess for a DOE that focuses λ_1 illumination into a round-shape spot and λ_2 illumination into a doughnut-shaped spot would be a DOE that fulfills at least half of its function. Because a binary-phase circular-symmetric DOE that best focuses λ_1 illumination into a spot is a Fresnel phase zone-plate, the initial population for GA was generated by adding Gaussian perturbations to a phase Fresnel zone plate designed for λ_1. The GA is first carried out by optimizing only for highest focusing efficiency of the λ_2 doughnut-shaped spot. The best solution of this first optimization step is then used as the parent for the full GA optimization. The initial population for the full GA optimization was generated similarly by adding Gaussian perturbations to the parent. The Gaussian perturbation used has a zero mean with a standard deviation ranging from $0.4\lambda_1$ to $50\lambda_1$. While larger standard deviation gives more diversity in the initial population, smaller standard deviations gives faster convergence and are useful when the initial population is close to the final solution.

4.3.3 The Fitness Function

The fitness function determines which children are preserved in evolution and is crucial for emphasizing particular aspects of the design and for balancing among various design merits. In order to optimize focal plane intensities in the round-shaped spot for λ_1 and ring-shaped spot for λ_2, intensity integrals over the λ_1 spot and λ_2 ring

are defined as followed,

$$E_1(\{\overrightarrow{r}, h\}) = \int_0^{\rho_1'} I(\rho', \lambda_1, \{\overrightarrow{r}, h\})\rho' d\rho' \tag{4.8}$$

$$E_2(\{\overrightarrow{r}, h\}) = \int_{\rho_2'}^{\rho_3'} I(\rho', \lambda_2, \{\overrightarrow{r}, h\})\rho' d\rho' \tag{4.9}$$

where ρ_1 is the nominal radius of the round spot at λ_1 and ρ_2 and ρ_3 are the nominal inner and outer radii of the ring-shaped spot at λ_2. In this work, the optimization was carried out with $\rho_1 = \lambda_1/\mathrm{NA}$, $\rho_2 = \lambda_2/\mathrm{NA}$, and $\rho_3 = 2\lambda_2/\mathrm{NA}$ to achieve focusing as tight as those allowed by conventional diffraction limit for both λ_1 and λ_2. In addition to focusing efficiencies, the quality of the node at λ_2 is also crucial for the operation of absorbance modulation. Therefore, a third term is added to the fitness function to adjust emphasis on the depth of the null at the center of the λ_2 ring-shaped spot. The typical fitness function that was used throughout most of the work looks as followed.

$$\begin{aligned} E(\{\overrightarrow{r}, h\}) &= -w_1 * E_1 - w_2 * E_2 + w_3 I(\rho' = 0, \lambda_2, \{\overrightarrow{r}, h\}) \\ &\equiv E_{total} = -w_1 * E_1 - w_2 * E_2 + w_3 * E_{node} \end{aligned} \tag{4.10}$$

where w_1, w_2, and w_3 are positive weights that allow relative emphasis of one term or another in the fitness function. Note the different signs in each term allow the negative terms to be maximized and the positive terms to be minimized when the overall fitness function is minimized by GA. The relative magnitudes of w_1, w_2, and w_3 determine which terms are weighted more heavily in the optimization process and will be described in more detail in the next section. In general, since the initial population was generated via slight variations to a phase zone plate at λ_1, the weight w_2 was chosen to be as high as 10 times w_1 in some cases. Although the higher the weight of w_3, the darker the null at λ_2, it should not be so large that it overwhelm the other terms in the fitness function. The weight of w_1 was set to 1 throughout this work.

4.4 Optimization results

As described in previous sections, a genetic algorithm (GA) is ideal for highly non-linear problems with a large set of parameters. However, evolution does not always guarantee best solution and can often get trapped in local optimal solutions. As a result, multiple runs of GA are often required to obtain solutions closer to global optima. Furthermore, different weight selections in the fitness function or functional changes in the definition of fitness functions gives additional flexibility in adjusting design emphasis. In this section, I will first show examples of HSQ dichromats designed with various weight settings in the fitness function, and then discuss how properties other than focusing efficiencies can be taken into account by adjusting functional forms of the fitness function with examples of PMMA dichormats [117]. HSQ and PMMA dichromats were investigated in parallel because HSQ was considered the ideal material, in which dichromats will be eventually manufactured, while PMMA was chosen for the first experimental demonstrations of dichromats for its ease of processing.

4.4.1 HSQ dichromats

The optimization technique based on GA was used to design dichromats with NA = 0.7 (80 zones) in HSQ. The refractive indices of HSQ, measured using a spectroscopic ellipsometer, were 1.4205 at $\lambda_1 = 405$ nm and 1.4100 at $\lambda_2 = 532$ nm. Because optical and other properties of HSQ are known to be sensitive to thermal processing applied before or after electron-beam exposures, it is important to note that these indices were obtained for HSQ samples baked in ambient with a furnace at 450 degrees Celsius for 2 hours.

Before going into details of how optimal weights were empirically determined, I'd like to revisit the fitness function definition. The first weight, w_1, in w=[w_1, w_2, w_3] maximizes focusing of $\lambda_1 = 405$ nm illumination into a round-shape spot, the second weight, w_2, maximizes focusing of $\lambda_2 = 532$ nm illumination into a ring-shape spot, and the third weight, w_3, minimizes intensity at the center of the λ_2 ring.

While higher efficiencies are in general desirable for DOE designs, for applications to absorbance modulation, it is also desirable to have as high E_2 to E_1 ratio as possible as the higher λ_2 to λ_1 peak intensity ratios, the more PSF compression at the output plane of the absorbance modulation layer (AML). w_3 is usually set to be orders of magnitude higher than w_1 to account for the difference in units and to enforce deeper nulls, as PSF compression is very sensitive to intensities at the null.

As described in section 4.3.2, an initial population for GA was obtained by first optimizing for only the second and third terms in Eq. 4.10, using radii of Fresnel zone plates as the initial population. This initial population, Initial 1, was optimized with $[w_1, w_2, w_3] = [0, 1, 10^4]$ and the resulting energy terms in the fitness function were $E_1 = 8.1483 \times 10^{18}$, $E_2 = 5.8802 \times 10^{16}$, $E_{node} = 4.0732 \times 10^{10}$, and $E_{total} = -5.8394 \times 10^{16}$. The set of radii is shown in Table 4.2 and was used as the initial population for the following two 80-zone designs discussed below. The phase step height, h, was fixed to give pi phase shift at $\lambda_1 = 405$ nm.

The first set of HSQ dichromat, Design 1, was obtained by picking the design with the lowest E_{total} among several GA runs with the same initial population and a weight of w=[1, 1, 10^4]. This design gave $E_1 = 5.7813 \times 10^{18}$, $E_2 = 7.97 \times 10^{16}$, $E_{node} = 4.5116 \times 10^{11}$, and $E_{total} = -5.8606 \times 10^{18}$. Although total energy of the fitness function is quite low, it can be identified that the fitness function is dominated by the E_1 term, as E_1 is almost two orders of magnitude larger than E_2 and $w_3 * E_{node}$. Comparing Design 1 with its initial population, both E_1 and E_{node} worsen, but E_2 became slightly higher than that in the initial population. This shows that the GA is indeed evaluating toward our design goal, an optic that maximizes both E_1, E_2 and minimizes E_{node}, but the emphasis on maximizing E_2 was not enough. When adjusting weights to w=[1 10 10^5] to supress E_1 dominance by an order, a more balanced solution, Design 2, was obtained. Design 2 gave $E_1 = 7.2744 \times 10^{17}$, $E_2 = 1.9868 \times 10^{17}$, $E_{node} = 1.0650 \times 10^{10}$, and $E_{total} = -2.7132 \times 10^{18}$. The focusing efficiency at λ_2 is improved at the expense of lower E_1, which is probably ok. Lower focusing efficiency at λ_1 means more λ_1 background and higher incident power required in the λ_1 beam. At the intensity level of interest here for absorbance modulation, increasing

Table 4.2: The radii of "Initial 1" optimized with $[w_1, w_2, w_3] = [0, 1, 10^4]$ from λ_1 Fresnel zone-plate radii. The resulting energy terms in the fitness function are $E_1 = 8.1483 \times 10^{18}$, $E_2 = 5.8802 \times 10^{16}$, $E_{node} = 4.0732 \times 10^{10}$, and $E_{total} = -5.8394 \times 10^{16}$.

Zone	Radius (nm)	Zone Width (nm)	Zone	Radius (nm)	Zone Width (nm)
1	3940.2		41	27235.1	352.9
2	5729.6	1789.3	42	27737.6	502.5
3	7019.9	1290.3	43	27940.1	202.5
4	8093.5	1073.7	44	28363.4	423.2
5	9110.1	1016.5	45	28685.6	322.2
6	10045.6	935.6	46	28952.7	267.1
7	10942.0	896.3	47	29351.0	398.3
8	11526.0	584.0	48	29717.0	366.0
9	12314.6	788.7	49	30084.9	367.9
10	13009.9	695.2	50	30403.9	318.9
11	13609.7	599.8	51	30684.7	280.8
12	14228.9	619.2	52	31021.9	337.2
13	14675.6	446.7	53	31339.9	318.0
14	15466.2	790.5	54	31728.3	388.4
15	15919.3	453.1	55	31930.6	202.3
16	16502.9	583.6	56	32417.8	487.2
17	17018.0	515.1	57	32810.6	392.8
18	17556.4	538.4	58	33031.3	220.7
19	18093.8	537.3	59	33429.8	398.5
20	18675.7	582.0	60	33765.1	335.3
21	19061.1	385.3	61	33978.7	213.6
22	19566.5	505.5	62	34281.5	302.8
23	20041.9	475.4	63	34542.6	261.1
24	20483.7	441.8	64	34904.5	361.9
25	20781.7	297.9	65	35284.3	379.8
26	21439.0	657.3	66	35523.0	238.7
27	21777.0	338.0	67	35813.6	290.7
28	22217.6	440.6	68	36120.5	306.9
29	22626.5	408.9	69	36521.0	400.4
30	22888.9	262.5	70	36723.4	202.5
31	23482.6	593.6	71	37107.2	383.8
32	24016.6	534.0	72	37322.1	214.9
33	24219.1	202.5	73	37524.6	202.5
34	24611.0	391.9	74	37914.8	390.2
35	24963.3	352.3	75	38195.9	281.1
36	25299.4	336.1	76	38506.0	310.1
37	25731.9	432.5	77	38960.6	454.7
38	26302.6	570.7	78	39201.1	240.5
39	26544.2	241.6	79	39414.5	213.4
40	26882.2	338.0	80	39681.7	267.2

power in λ_1 is usually not a limitation. It is the power in λ_2 that usually limits the performance, *i.e.* the highest λ_2 to λ_1 intensity ratio, for absorbance modulation. Absorbance modulation is also less sensitive to background λ_1 illumination compared to background at λ_2. The radii of the designs are listed in Table 4.3 and 4.4.

In addition to adjusting weights for optimization, different local optima and many times better solutions can be reached by simply re-running GA many times. There is a difference between the convergence of each GA run and the convergence to a global optimal solution. Each GA run ends either when the convergence criteria are met or when maximum number of iterations, 100, is reached. With the GA settings described above, most GA runs converge within 50 iterations. An example of the GA iteration output is shown in Table 4.5. "Generation" represents the number of iterations, "f-count" is the cumulative number of evaluations of the fitness function, "Best f(x)" is the lowest fitness function value achieved in the population, "Mean f(x)" is the mean fitness value of the population, "Stall Generations" is the number of consecutive iterations without improvement. As the population evolves, "Best f(x)" decreases rapidly and while "Mean f(x)" gradually follows. Convergence is met when "Mean f(x)" approaches "Best f(x)."

In every GA run, the solution converges to a local optimum, but it was found that solutions don't necessary converge to a single design even with the same initial population and weight vector. (Note: In fact, the initial populations of each GA run are not exactly the same because of the Gaussian perturbations added to the set of initial populations.) There is a trade-off when setting the allowed level of mutation of the GA. When mutation is high, GA is less likely to be trapped in local optima, but the solution converges much slower and may require a second refining search step to accurately achieve the true optimum. Global optimization is difficult to ensure with the structure of this design problem unless infinite paths of evolution are taken. Part of the reason may be because the adaptive mutation method chosen here tends to send evolution towards its initial path to increase speed of convergence within each round. The "adaptive feasible mutation function" randomly generates directions that are adaptive with respect to the last successful or unsuccessful generation and

Table 4.3: The radii of "Design 1" optimized with $[w_1, w_2, w_3] = [1, 1, 10^4]$.

Zone	Radius (nm)	Zone Width (nm)	Zone	Radius (nm)	Zone Width (nm)
1	3893.3		41	27267.1	418.7
2	5647.7	1754.4	42	27695.5	428.4
3	6967.2	1319.5	43	27941.1	245.6
4	7700.6	733.4	44	28358.1	417.0
5	9136.4	1435.8	45	28655.9	297.8
6	10036.3	900.0	46	29076.0	420.1
7	10769.9	733.6	47	29390.4	314.5
8	11360.1	590.2	48	29668.2	277.8
9	12717.0	1356.9	49	30091.6	423.3
10	13461.8	744.8	50	30398.4	306.8
11	13672.4	210.6	51	30655.3	256.9
12	13914.1	241.7	52	31019.4	364.1
13	14675.7	761.5	53	31344.9	325.5
14	15440.6	764.9	54	31727.9	383.0
15	15883.2	442.6	55	32002.2	274.4
16	16480.2	597.0	56	32441.3	439.1
17	17038.1	557.9	57	32883.4	442.1
18	17633.7	595.6	58	33092.0	208.6
19	18151.3	517.6	59	33484.4	392.4
20	18676.6	525.2	60	33752.1	267.6
21	19095.7	419.1	61	33970.8	218.8
22	19627.5	531.8	62	34326.4	355.5
23	20070.2	442.7	63	34551.6	225.2
24	20410.3	340.1	64	34922.5	370.9
25	20827.6	417.3	65	35339.4	417.0
26	21489.1	661.6	66	35582.4	243.0
27	21808.2	319.1	67	35846.1	263.7
28	22194.2	385.9	68	36090.9	244.8
29	22578.5	384.3	69	36535.9	445.0
30	22837.5	259.0	70	36759.9	224.0
31	23551.0	713.4	71	37115.9	356.0
32	23865.5	314.5	72	37318.0	202.2
33	24106.9	241.5	73	37539.0	220.9
34	24496.8	389.9	74	37901.4	362.4
35	24994.5	497.7	75	38126.9	225.5
36	25302.6	308.1	76	38379.6	252.7
37	25744.4	441.8	77	38924.9	545.3
38	26316.3	571.8	78	39192.4	267.5
39	26553.8	237.5	79	39395.0	202.6
40	26848.4	294.6	80	39667.2	272.2

Table 4.4: The radii of "Design 2" optimized with $[w_1, w_2, w_3] = [1, 10, 10^5]$.

Zone	Radius (nm)	Zone Width (nm)	Zone	Radius (nm)	Zone Width (nm)
1	14617.4		41	4127333.5	235.1
2	25938.7	1321.3	42	4227839.0	505.5
3	37597.3	1658.6	43	4328043.9	204.9
4	48398.0	800.6	44	4428438.7	394.8
5	59046.4	648.5	45	4528641.4	202.7
6	69661.8	615.4	46	4628843.9	202.5
7	79865.2	203.3	47	4729140.7	296.7
8	811434.2	1569.1	48	4829422.3	281.6
9	912944.3	1510.1	49	4930252.8	830.5
10	1013147.4	203.1	50	5030457.5	204.8
11	1113379.1	231.7	51	5131310.5	853.0
12	1214731.1	1352.0	52	5231544.4	233.8
13	1314989.9	258.8	53	5331749.0	204.6
14	1415703.8	713.9	54	5431951.5	202.6
15	1516104.4	400.6	55	5532372.9	421.4
16	1616594.3	490.0	56	5632577.4	204.5
17	1717077.6	483.2	57	5732779.9	202.5
18	1817904.6	827.0	58	5833227.2	447.2
19	1918107.3	202.7	59	5933459.7	232.6
20	2018733.4	626.2	60	6033699.3	239.6
21	2118936.2	202.8	61	6134129.8	430.6
22	2219361.6	425.3	62	6234337.6	207.8
23	2320008.7	647.1	63	6334540.7	203.1
24	2420730.0	721.3	64	6435230.3	689.6
25	2520936.8	206.8	65	6535432.9	202.6
26	2621139.5	202.7	66	6635636.0	203.1
27	2721344.2	204.7	67	6735850.1	214.1
28	2822758.5	1414.3	68	6836052.5	202.5
29	2922961.0	202.5	69	6936256.1	203.6
30	3023163.7	202.7	70	7036983.1	727.1
31	3123575.9	412.2	71	7137185.6	202.5
32	3223779.6	203.6	72	7237388.0	202.4
33	3323986.9	207.3	73	7337590.7	202.7
34	3424395.7	408.8	74	7437874.5	283.8
35	3524598.4	202.7	75	7538241.2	366.6
36	3625438.9	840.5	76	7638443.7	202.5
37	3725934.1	495.1	77	7738646.2	202.5
38	3826136.6	202.5	78	7838924.8	278.6
39	3926889.9	753.3	79	7939156.8	232.0
40	4027098.3	208.5	80	8039681.6	524.8

Table 4.5: An example of the Genetic Algorithm (GA) iteration output is shown in the table. "Generation" represents the number of iterations, "f-count" is the cumulative number of evaluations of the fitness function, "Best f(x)" is the lowest fitness function value achieved in the population, "Mean f(x)" is the mean fitness value of the population, "Stall Generations" is the number of consecutive iterations without improvement. As the population evolves, "Best f(x)" decreases rapidly and while "Mean f(x)" gradually follows. Convergence is met when "Mean f(x)" approaches "Best f(x)." The energies printed at the end was not part of the GA output, but a separate function called to compute energies for the best solution returned by GA. The simulation time of this GA optimization was about 5 hours.

Generation	f-count	Best f(x)	Mean f(x)	Stall Generations
1	50	4.586e+022	6.398e+022	0
2	75	1.076e+022	6.288e+022	0
3	100	1.076e+022	6.301e+022	1
4	125	2.314e+021	3.914e+022	0
5	150	2.314e+021	3.082e+022	1
6	175	2.314e+021	1.818e+022	2
7	200	1.964e+021	1.218e+022	0
8	225	1.964e+021	8.603e+021	1
9	250	2.481e+020	6.262e+021	0
10	275	2.481e+020	4.128e+021	1
11	300	8.653e+019	4.038e+021	0
12	325	8.653e+019	3.306e+021	1
13	350	8.653e+019	2.338e+021	2
14	375	1.049e+019	1.734e+021	0
15	400	2.318e+018	1.59e+021	0
16	425	2.318e+018	2.389e+021	1
17	450	2.318e+018	1.494e+021	2
18	475	2.318e+018	8.297e+020	3
19	500	2.318e+018	6.803e+020	4
20	525	2.318e+018	4.589e+020	5
21	550	2.318e+018	2.623e+020	6
22	575	-9.211e+017	6.722e+019	0
23	600	-1.017e+018	1.831e+019	0
24	625	-1.429e+018	2.989e+019	0
25	650	-1.474e+018	5.606e+019	0
26	675	-1.474e+018	3.578e+020	1
27	700	-1.474e+018	1.641e+020	2
28	725	-1.474e+018	1.14e+020	3
29	750	-1.474e+018	8.395e+019	4
30	775	-1.474e+018	5.063e+019	5

Table 4.6: Table 4.5 continued.

Generation	f-count	Best f(x)	Mean f(x)	Stall Generations
31	800	-1.7e+018	3.162e+019	0
32	825	-1.7e+018	2.203e+019	1
33	850	-1.7e+018	1.776e+019	2
34	875	-1.801e+018	5.959e+017	0
35	900	-1.801e+018	4.479e+017	1
36	925	-1.801e+018	2.918e+017	2
37	950	-1.801e+018	7.415e+017	3
38	975	-1.801e+018	2.413e+018	4
39	1000	-1.801e+018	6.246e+017	5
40	1025	-1.805e+018	-2.448e+017	0
41	1050	-1.805e+018	-1.26e+018	1
42	1075	-1.805e+018	-1.419e+018	2
43	1100	-1.805e+018	-1.499e+018	3
44	1125	-1.805e+018	-1.664e+018	4
45	1150	-1.805e+018	-1.65e+018	5
46	1175	-1.805e+018	-1.669e+018	6
47	1200	-1.805e+018	-1.662e+018	7
48	1225	-1.805e+018	-1.663e+018	8
49	1250	-1.805e+018	-1.635e+018	9
50	1275	-1.805e+018	-1.685e+018	10
51	1300	-1.805e+018	-1.728e+018	11

Optimization terminated: average change in the fitness value less than options.TolFun.

Energy = -1.8047e+018
E1 = 4.1859e+017
E2 = 1.3863e+017
Enod = 1.5783e+009

Table 4.7: A summery of energies and the corresponding w vector used, for Initail 1, Design 1, Design 2, and Initial 2.

	w	E_{total}	E_1	E_2	E_{node}
Initial 1	[0, 1, 10^4]	-5.8394E+16	8.1483E+18	5.8802E+16	4.0732E+10
Design 1	[1 1 10^4]	-5.8606E+18	5.7813E+18	7.97E+16	4.5116E+11
Design 2	[1 10 10^5]	-2.7132E+18	7.2744E+17	1.9868E+17	1.0650E+10
Initial 2	[0, 1, 10^4]	-2.8888E+17	2.4292E+17	2.8975E+17	8.7362E+10

adjusts step sizes to ensure new generations remain in the feasible region bounded by the inequality constraints [116]. Another obvious reason for limited convergence is due to the small size of population and number of generations. Most work that applies GA for source mask optimization requires at many as 500-5000 iteration to converge [118, 119, 120]. By simply running GA more times for the initial population optimization with w= [w_1, w_2, w_3] = [0, 1, 10^4], a second initial population, Initial 2, is achieved. The energy terms of Initial 2 was $E_1 = 2.4292 \times 10^{17}$, $E_2 = 2.8975 \times 10^{17}$, $E_{node} = 8.7362 \times 10^{10}$, and $E_{total} = -2.8888 \times 10^{17}$, significantly outperforming the previous initial population, Initial 1, in terms of E_2 and E_{total}. A table summarizing the four different designs are shown below in Table 4.9 and the radii of Initial 2 is shown in Table 4.8.

Better initial populations help GA achieve superior solutions. Notice that E_1 and E_2 are more balanced in Initial 2 compared to Initial 1. This makes it easier to retain that balance, *i.e.* to prevent E_1 term from dominating the fitness function. The following designs were selected from the solutions of 50 GA runs from Initial 2 with w = [w_1, w_2, w_3]=[1, 10, 10^5], same as Design 2. Among the 50 runs, Design 3 yielded the lowest total energy and highest E_2, while Design 4 yielded the lowest E_{node}. Note that although both designs showed lower total energy compared to Design 2, their E_1 values are not as high as that for Design 2. Among the 50 runs, there was indeed one solution (Design 5) with lower total energy, higher E_1, and higher E_2 compared to Design2, but its E_{node} value was higher. The point of this study was to show that this optimization problem intrinsically exhibits many local optima because the fitness function can be locally improved by optimizing for either term in the fitness function. As a result, multiple GA runs are required to search for a more globally

Table 4.8: The radii of "Inital 2" optimized with $[w_1, w_2, w_3] = [1, 10, 10^5]$ from the Fresnel zone-plate radii.

Zone	Radius (nm)	Zone Width (nm)	Zone	Radius (nm)	Zone Width (nm)
1	4025.3		41	27073.2	211.8
2	5706.1	1680.8	42	27648.8	575.6
3	7295.4	1589.3	43	28047.3	398.5
4	8038.1	742.8	44	28422.1	374.8
5	9415.9	1377.8	45	28626.2	204.1
6	10539.9	1123.9	46	29635.4	1009.2
7	10784.1	244.2	47	30061.9	426.5
8	11250.9	466.8	48	30274.3	212.4
9	12301.1	1050.1	49	30476.8	202.5
10	12633.0	331.9	50	30682.1	205.3
11	14250.0	1617.0	51	30885.1	203.0
12	14473.7	223.8	52	31087.6	202.6
13	14874.0	400.3	53	31453.4	365.8
14	15174.5	300.5	54	31679.3	225.9
15	15382.9	208.4	55	31955.7	276.4
16	15587.4	204.5	56	32741.6	785.8
17	15972.0	384.7	57	32947.2	205.7
18	18058.8	2086.8	58	33281.4	334.2
19	18927.6	868.7	59	33634.2	352.8
20	19151.8	224.2	60	33853.9	219.7
21	19473.6	321.8	61	34058.6	204.7
22	19851.5	377.9	62	34260.8	202.2
23	20058.8	207.3	63	34463.8	203.0
24	20535.8	477.1	64	34675.1	211.4
25	20754.6	218.8	65	35035.2	360.1
26	21133.0	378.4	66	35537.4	502.2
27	21358.6	225.6	67	35742.2	204.8
28	22393.7	1035.1	68	35947.0	204.7
29	22820.3	426.6	69	36153.6	206.7
30	23445.8	625.5	70	36385.3	231.7
31	23658.1	212.3	71	36736.1	350.8
32	24137.1	479.0	72	37591.1	855.0
33	24527.3	390.2	73	37794.4	203.3
34	24774.8	247.5	74	38063.6	269.1
35	25276.2	501.4	75	38266.2	202.6
36	25959.6	683.3	76	38468.7	202.5
37	26172.2	212.6	77	38673.7	205.0
38	26414.3	242.1	78	38880.5	206.8
39	26658.9	244.6	79	39385.6	505.1
40	26861.4	202.5	80	39589.6	204.0

Table 4.9: Summery of energies and the corresponding w vector used, for Design 3-7.

	w	E_{total}	E_1	E_2	E_{node}
Design 3	[1, 10, 10^5]	-5.32E+18	2.33E+17	5.10E+17	3.55E+10
Design 4	[1 10 10^5]	-2.74E+18	5.94E+17	2.15E+17	2.73E+09
Design 5	[1 10 10^5]	-3.72E+18	7.53E+17	2.97E+17	3.01E+10
Design 6	[1, 1, 10^7]	-1.08E+18	7.94E+17	4.88E+17	2.01E+10
Design 7	[1, 1, 10^7]	-2.38E+17	1.19E+17	1.21E+17	1.94E+08

optimal solution and even with as many as 50 GA runs, we can only conclude that we've approached the global optimum.

Finally, in order to test how far the GA can suppress the node intensity, a weight function that highly emphasize the third term of the fitness function, $w = [w_1, w_2, w_3]=[1, 1, 10^7]$, was chosen for another 50 GA runs with an initial population generated from Initial 2. The best solution, Design 6, yielded E_1, E_2, and E_{node} values similar to those of Design 3-5, disregarding the fact the w is quite different for the two. (Note that total energy values cannot be directly compared in this case because the weight functions are different.) Design 7 gave an E_{node} as low as 1.94×10^8, but as will become obvious in the next section, true node quality requires FDTD simulations to evaluate. Therefore, w_3 was not highly emphasized for most cases as the E_{node} obtained here may not be as accurate representation of how deep the node is.

The radii of Design 3 to Design 7 shown in the Appendix B.

4.4.2 PMMA Dichromats

Dichromats fabricated in PMMA were first investigated for its ease of processing as an electron-beam resist. PMMA optics were optically characterized and used to conduct absorbance modulation experiments, and so were most thoroughly investigated. The work on PMMA dichromat designs were led mainly by Dr. Rajesh Menon with significant contributions from Paul Rogge, who worked at MIT as a summer student of the National Science Foundation (NSF) REU program.

The refractive indices of PMMA as measured using a spectroscopic ellipsometer were 1.501 at $\lambda_1 = 405$ nm and 1.487 at $\lambda_2 = 532$ nm. It was determined empirically

that w=[w_1, w_2, w_3]=[1, 10, 10^4] was appropriate for the following PMMA designs. Only 40 and 80-zone designs are discussed in this section and the zone heights, h, were not fixed to a certain value. The dichromats that were actually fabricated have NAs ranging from 0.22 to 0.83 and all shared the same h, except for the NA=0.7 dichromat, which was first fabricated on a different substrate.

The transmission functions of the dichromats after 25 generations of GA are shown in Fig. 4-2(a) for an optimal 40-zone design and in Fig. 4-2(c) for the 80-zone design. The radii of the zones and the height, h, are listed in a publication [117]. Radial cross sections of the intensity distribution through the centers of the focal spots are plotted in Figs. 4-2(b) and 4-2(d) for the designs in Figs. 4-2(a) and 4-2(c), respectively. Although the design was performed using scalar diffraction theory, as seen in the expression of the fitness function, a rigorous full-vector solution of Maxwell's equations using the finite-difference time domain (FDTD) method [121] was used to simulate focal spots of the designs to confirm that the dichromats indeed focuse λ_1 to a round bright spot and λ_2 to a ring-shaped spot. A circular polarized uniform plane wave was assumed for the FDTD simulations. Figs. 4-2(b) and 4-2(d) show the FDTD data plotted with x's and the scalar data plotted with solid curves. Although features on the order of $0.5\lambda_1$ are present in the dichromats and the NA of the 80-zone dichrmats is as high as 0.7, The general shape of point-spread function (PSF) obtained with scalar theory agrees with that obtained with FDTD. The focal node at λ_2 is formed as a result of destructive interference from the concentric zones; no phase singularity is imposed. Note that the quality of the node, *i.e.* λ_2 intensity at $\rho = 0$, worsens as the dichromat NA increases according to FDTD simulations, but the same behavior is not observed with scalar theory simulations. This implies that scalar diffraction theory may not accurately represent how deep the λ_2 node is, especially for high NA dichromats. Therefore, optimization with heavy emphasis on minimizing λ_2 node intensities obtained by scalar diffraction theory, *i.e.* with a large w_3 value and the current fitness function, may not be very meaningful. In order to accurately optimize "darkness" of the node, a fitness function that incorporates FDTD simulations would be required.

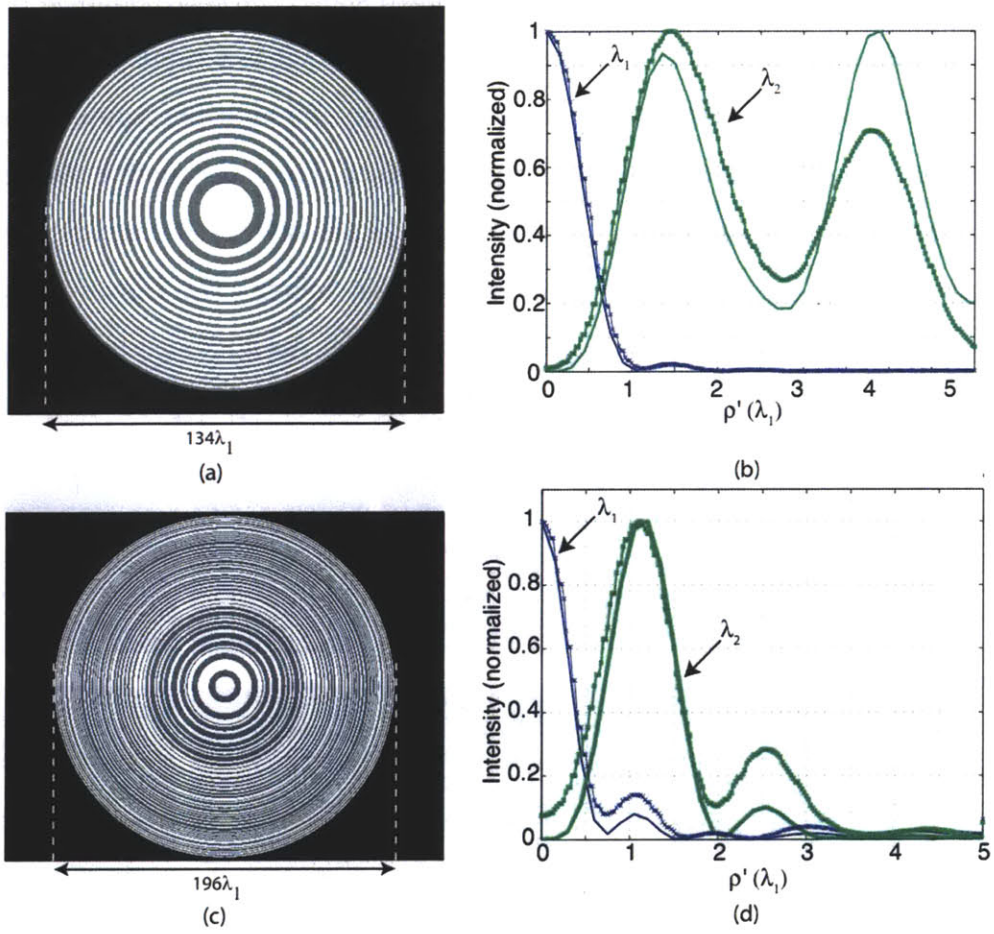

Figure 4-2: Dichromat, all lengths are in units of λ_1. Transmission functions of dichromats of (a) NA=0.56 and (c) NA=0.7, both with focal length=$100\lambda_1$, λ_1 = 405 nm, λ_2=532 nm, n(λ_1)=1.501, and n(λ_2)=1.487. The remaining optimization parameters are described in the text. Within the aperture of the dichromat, the gray rings are phase shifted with respect to the white ones; their relative height difference being (a) $0.98\lambda_1$ and (c) $0.74\lambda_1$. Outside the dichromat is opaque. (b) and (d) Radial intensity distributions of the focal spots. The data calculated using FDTD are plotted with x's while the scalar data are plotted as solid curves.

Defocus

Figure 4-3(a) shows the focused spots calculated using FDTD for various values of z. The depth-of-focus (DOF) of the lens is defined as the range where the ring at λ_2 is well-defined and is approximately $\lambda_1/2$. Fig. 4-3(b) shows the corresponding data for λ_1. Notice that the intensity in the central node at λ_2 increases rapidly with defocus. This is characteristic of destructive interference from the multiple zones. When the appropriate path-difference condition is not met, there is no destructive interference. This is problematic for patterning, where a large depth-offocus is desired. Attempts to enlarge the DOF of dichromats by design include defining the fitness function to include a DOF term similar to the slope term defined for Aerial images for lithography patterns [119]. Further investigation is required on this design aspect.

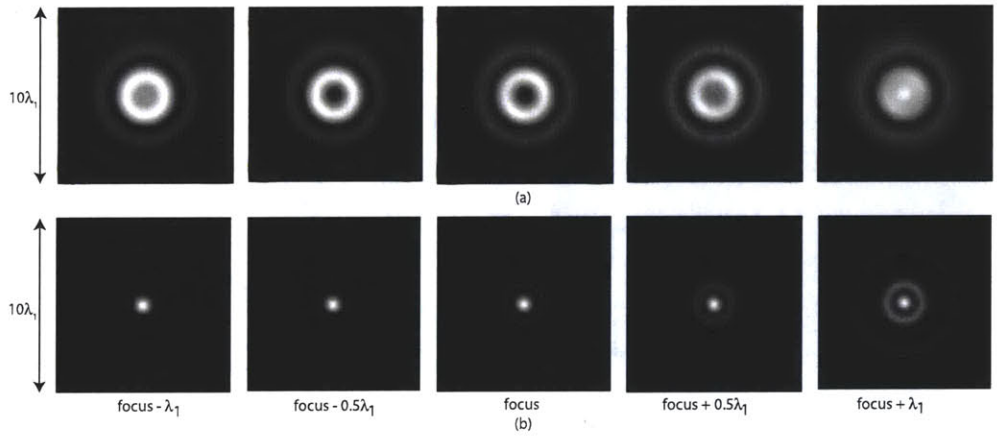

Figure 4-3: Intensity distributions calculated using FDTD of the dichromat shown in Fig. 4-2(c) at various transverse planes near the focus of the dichromat for (a) λ_2 and (b) λ_1.

Tolerance to Fabrication Errors

Because the dichromat would be fabricated using a scanning-electron-beam lithography process, it is important to understand the sensitivity of the focusing characteristics to fabrication errors. Considering only errors in the radii of the zones and the height of the zones, their effects can be simulated by adding randomly generated

errors to the zone radii and the phase height as followed:

$$\overrightarrow{r} = \{r_1, r_2, \ldots, r_M\} \rightarrow \quad \widetilde{\overrightarrow{r}} = \{r_1 + \widetilde{\delta_{r_1}}, r_2 + \widetilde{\delta_{r_2}}, \ldots, r_M + \widetilde{\delta_{r_M}}\} \tag{4.11}$$

$$h \rightarrow \qquad \widetilde{h} = h + \widetilde{\delta_h}, \tag{4.12}$$

where $\widetilde{\delta_r}$ and $\widetilde{\delta_h}$ are randomly generated from two normal distributions of zero mean and standard deviations, σ_r and σ_h, respectively. To quantify the effect of the error, we calculated the distribution of the focusing efficiencies of the error-prone dichromats at the two wavelengths and then calculated their corresponding standard deviations.

The focusing efficiencies are defined as:

$$\eta_1(\{\overrightarrow{r}, h\}) = \frac{\int_0^{\rho_1'} I(\rho', \lambda_1, \{\overrightarrow{r}, h\})\rho' d\rho'}{\int_0^{\infty} I(\rho', \lambda_1, \{\overrightarrow{r}, h\})\rho' d\rho'}, \tag{4.13}$$

$$\eta_2(\{\overrightarrow{r}, h\}) = \frac{\int_{\rho_2'}^{\rho_3'} I(\rho', \lambda_2, \{\overrightarrow{r}, h\})\rho' d\rho'}{\int_0^{\infty} I(\rho', \lambda_2, \{\overrightarrow{r}, h\})\rho' d\rho'}. \tag{4.14}$$

The robustness of the dichromat was investigated by calculating the standard deviations of the focusing efficiencies, $\sigma\{\eta_1(\{\overrightarrow{r}, h\})\}$ and $\sigma\{\eta_2(\{\overrightarrow{r}, h\})\}$, as a function of σ_r and σ_h. The results for the original dichromat shows that the variations in η_1 and in η_2 are below 1% and 4%, respectively, even for large fabrication errors (σ_r=0.4λ_1=160nm and σ_h=0.04λ_1=16nm). To design a more robust dichromat, we defined a new fitness function as follows:

$$E_r(\{\overrightarrow{r}, h\}) = \mu\{E(\{\widetilde{\overrightarrow{r}}, \widetilde{h}\})\} + \sigma\{E(\{\widetilde{\overrightarrow{r}}, \widetilde{h}\})\}, \tag{4.15}$$

where $\mu\{\}$ and $\sigma\{\}$ represent mean and standard deviations, respectively. This fitness function was used to design a new 80-zone dichromat with the same parameters as described earlier. Standard deviations, σ_r=$\lambda_1/8$ and σ_h=$\lambda_1/40$, were used during the optimization. The initial population for the genetic algorithm was derived by adding small perturbations to the original 80-zone dichromat. After five generations, the variation in the focusing efficiencies at the two wavelengths are somewhat suppressed in the new design for both wavelengths. Further improvements can be expected with

more generations as well as improved fitness functions. All efficiencies were calculated using the scalar theory.

4.4.3 Summary and Future Work

In summary, we developed and verified a methodology for circular-symmetric diffractive-optical-element design for specific illumination patterns at multiple wavelengths. We obtained designs of dichromats that focuses $\lambda_1 = 405$ nm plane-wave illumination into a round-shape spot, and $\lambda_2 = 532$ nm plate-wave illumination into a doughnut-shaped spot. We found that a two-step optimization process, starting from conventional λ_1 zone plates, optimizing first for the λ_2 illumination, and then for the entire design, gives reasonable focusing efficiencies for the desired patterns. The effects of the relative weights of various terms in the fitness function was studied and tuned to guide evolution of the optimization algorithm. The dichromats obtained show reasonable efficiencies and a node quality at least below 10% of the peak values. The use of scalar diffraction theory for evaluating the fitness function, which tends to overestimate the darkness of the node, prevented the algorithm from accurately determining the quality of the node, especially for high NA designs. An FDTD solver would be ideal for accuracy, but terrible for throughput of the optimization algorithm. The effect of defocus and tolerance to fabrication errors were studied. Attempts to improve depth of focus and tolerance to fabrication errors by adding additional terms in the fitness function resulted in some improvements. However, only zone width and zone height errors were considered in this case, while other fabrication errors, such as astigmatism, can also affect the output of dichromats. Finally, extending (or reducing) this optimization technique to design single wavelength optics with tighter PSFs at the expense of higher side-lobes can also be useful for future applications with absorbance modulation.

The GA optimization codes are documented in Appendix B.

4.5 PMMA Dichromat Fabrication

Phase-only diffractive lenses with circular symmetry can be readily fabricated in a dielectric material using planar processes with scanning-electron beam lithography, enabling large arrays with high optical uniformity [122]. As shown in Eq. 4.2 and Fig. 4-1, the phase shift is achieved via a height difference between alternate zones, which can be accurately controlled by spinning of the electron-beam resist.

PMMA were chosen to be the phase-shift material for a first demonstration of dichromat designs. PMMA, polymethylmethacrylate, with a molecular weight of 950K is a commonly used positive-tone e-beam resist. PMMA provides high contrast, structural rigidity, transparency at visible wavelength, and high robustness to environmental conditions (especially compared to HSQ), such as pre-exposure baking conditions and conductive layer selection. One major drawback of PMMA is that it dissolves in most cleaning solvents used in the lab, for example acetone and methanol, so cleaning PMMA optics becomes difficult. In addition, PMMA is relatively soft, and can be easily scratched when placed in contact with samples by accident. Hard-baked HSQ optics will have fewer of these problems because HSQ becomes silicon dioxide after hard bake.

An illustration of the dichromat design patterns are shown in Fig. 4-4. The zone radii and the phase shift of the PMMA dichromats that were fabricated are shown in Table 4.10, Table 4.11, and Table 4.13. The corresponding energy values are listed in the table captions.

These patterns are converted into writing patterns described in an ASCII file for the RAITH-150 scanning-electron-beam-lithography (SEBL) system. The writing pattern consists of single-pixel concentric circular lines with 18nm line-spacing that fills the desired writing area. Each line consists of 1000 points (maximum for the tool) and runs through a full circle with a random start point to avoid cumulative extra exposures at starting points. The code for ASCII file generation is documented in Appendix B.

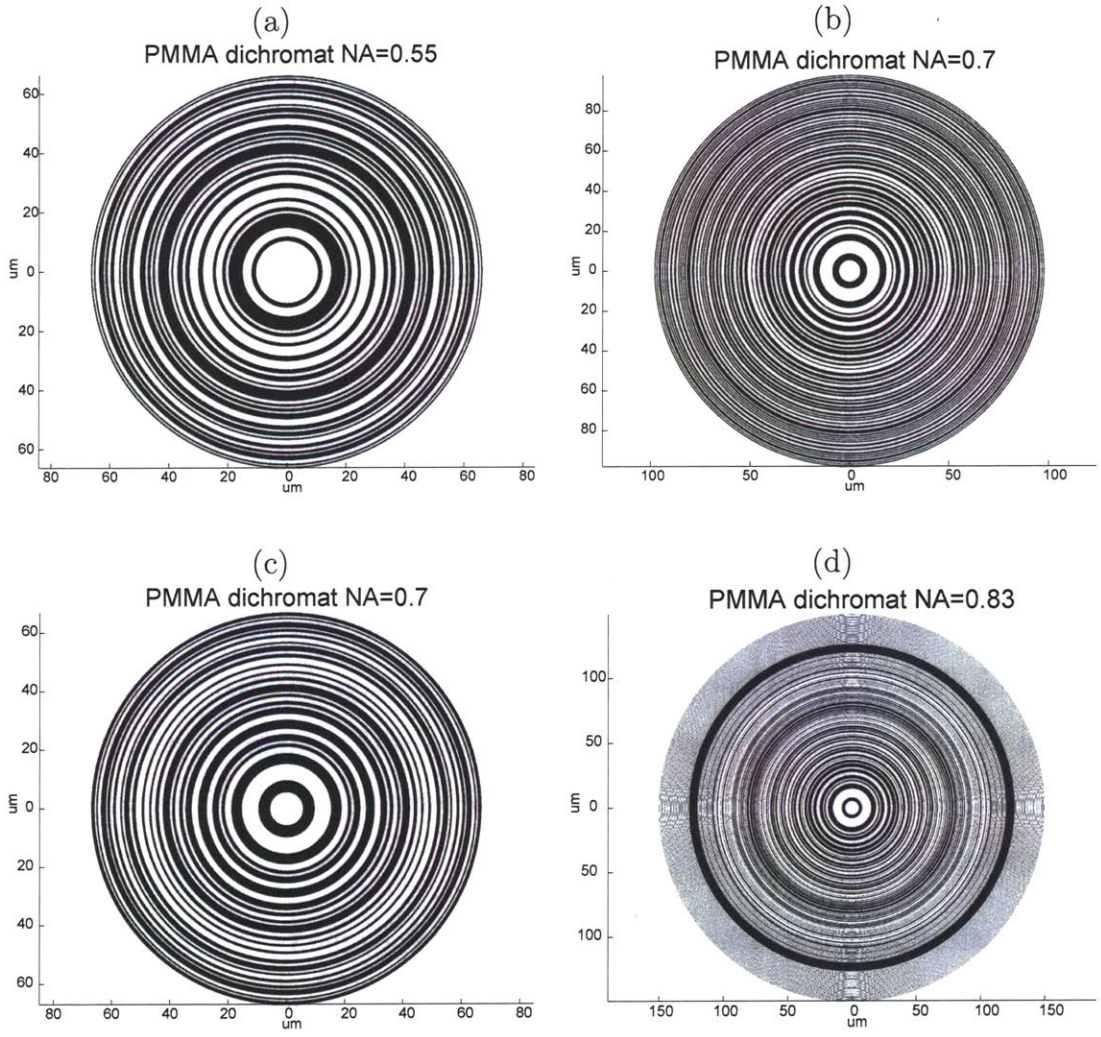

Figure 4-4: Designs for the fabricated dichromats in PMMA. All dichromats except the one in (b) were designed to have a π phase shift, which corresponds to a 400 nm phase step. The phase step in (b) was 338 nm. In the case for PMMA optics, black zones represent exposed lines, while white zones represent remaining PMMA lines. The PSFs for (a)(b)(d) were experimentally characterized below. (a) NA = 0.55 (40 zone) (b) NA = 0.7 (80 zone) (c) NA = 0.7 (80 zone) (d) NA = 0.83 (160 zone)

Table 4.10: The radii of the PMMA dichromat with NA=0.55, which we fabricated and characterized its point-spread function (PSF.) (phase step = 400 nm) E_1 = 1.5545E+17, E_2 = 1.2823E+17, E_{node} = 2.8520E+14.

Zone	Radius (nm)	Zone Width (nm)	Zone	Radius (nm)	Zone Width (nm)
1	4207.2		21	18465.7	298.5
2	4804.6	597.4	22	18853.0	387.3
3	6001.2	1196.6	23	19143.1	290.1
4	7895.3	1894.1	24	19887.9	744.8
5	8282.4	387.1	25	20765.3	877.4
6	8750.5	468.1	26	21308.0	542.8
7	9604.7	854.2	27	21617.6	309.6
8	10034.4	429.7	28	22194.6	576.9
9	11353.8	1319.4	29	22486.5	292.0
10	12030.2	676.4	30	22691.8	205.3
11	13156.5	1126.3	31	23385.8	694.0
12	13717.2	560.6	32	23849.0	463.3
13	14141.0	423.9	33	24247.8	398.8
14	14814.7	673.6	34	24587.4	339.6
15	15379.7	565.0	35	24787.5	200.1
16	15606.4	226.8	36	25427.9	640.4
17	15976.8	370.4	37	25918.4	490.6
18	17421.1	1444.3	38	26131.7	213.2
19	17881.9	460.8	39	26332.8	201.2
20	18167.2	285.2	40	26533.0	200.2

Table 4.11: The radii of the PMMA dichromat with NA=0.7, which we fabricated and characterized its point-spread function (PSF.) (phase step = 338 nm) E_1 = 7.9899E+17, E_2 = 1.4481E+18, E_{node} = 3.7323E+15.

Zone	Radius (nm)	Zone Width (nm)	Zone	Radius (nm)	Zone Width (nm)
1	2177.8		41	27191.9	354.9
2	3510.6	1332.8	42	27583.6	391.8
3	6079.2	2568.6	43	27799.4	215.7
4	7411.0	1331.9	44	28199.4	400.0
5	8487.9	1076.8	45	28718.8	519.5
6	8797.2	309.3	46	29106.2	387.4
7	9331.9	534.8	47	29368.0	261.8
8	10243.8	911.9	48	29619.9	251.9
9	11068.7	824.9	49	29829.8	209.9
10	12121.7	1053.0	50	30081.6	251.8
11	12919.4	797.7	51	30527.2	445.6
12	13867.0	947.6	52	30776.8	249.6
13	14273.0	406.0	53	31209.7	432.9
14	14755.8	482.8	54	31619.4	409.7
15	15261.2	505.4	55	31995.0	375.6
16	15669.2	408.0	56	32666.8	671.8
17	16010.0	340.8	57	32915.4	248.6
18	16845.9	835.9	58	33197.8	282.4
19	17411.6	565.7	59	33397.8	200.0
20	18095.9	684.3	60	33597.8	200.0
21	18603.4	507.5	61	33797.8	200.0
22	18972.6	369.2	62	33997.8	200.0
23	19553.6	580.9	63	34205.1	207.3
24	19764.0	210.4	64	34420.5	215.4
25	20507.3	743.3	65	34828.2	407.8
26	21124.8	617.4	66	35228.5	400.3
27	21563.3	438.5	67	35651.7	423.2
28	22139.1	575.8	68	35999.7	348.0
29	22456.4	317.4	69	36441.9	442.2
30	22791.1	334.7	70	36771.2	329.3
31	23351.0	559.9	71	36979.5	208.2
32	23864.2	513.2	72	37213.8	234.3
33	24202.8	338.6	73	37413.8	200.0
34	24584.8	382.0	74	37613.8	200.0
35	24813.0	228.3	75	37880.1	266.3
36	25195.6	382.5	76	38344.0	463.9
37	25804.1	608.5	77	38583.2	239.2
38	26104.6	300.5	78	38783.2	200.0
39	26407.5	303.0	79	38983.2	200.0
40	26837.0	429.4	80	39191.8	208.6

Table 4.12: The radii of the PMMA dichromat with NA=0.7 that was fabricated on the same sample as those in Table 4.10 and Table 4.13. (phase step = 400 nm) E_1 = 8.6103E+17, E_2 = 1.2505E+18, E_{node} = 6.2623E+15.

Zone	Radius (nm)	Zone Width (nm)	Zone	Radius (nm)	Zone Width (nm)
1	2271.2		41	27328.8	542.9
2	3877.6	1606.4	42	27530.3	201.6
3	6132.5	2254.9	43	27833.6	303.3
4	7564.2	1431.6	44	28432.1	598.5
5	8458.2	894.1	45	28632.1	200.0
6	8903.2	445.0	46	28989.2	357.2
7	9303.2	400.0	47	29311.0	321.8
8	10157.0	853.8	48	29512.6	201.6
9	10928.4	771.4	49	29814.9	302.2
10	12120.0	1191.6	50	30182.4	367.5
11	12981.5	861.4	51	30405.6	223.3
12	13834.2	852.7	52	30831.1	425.5
13	14222.4	388.2	53	31243.0	411.9
14	14699.4	477.0	54	31648.9	405.9
15	15287.6	588.2	55	31989.6	340.6
16	15650.0	362.4	56	32659.0	669.4
17	15999.9	349.9	57	32859.9	200.9
18	16974.5	974.6	58	33157.4	297.4
19	17522.1	547.6	59	33357.4	200.0
20	17971.2	449.1	60	33646.8	289.4
21	18542.1	570.9	61	33847.0	200.2
22	18822.7	280.6	62	34047.1	200.1
23	19513.8	691.1	63	34248.2	201.1
24	19828.4	314.6	64	34558.1	309.9
25	20670.4	841.9	65	34793.3	235.2
26	21137.6	467.2	66	35004.7	211.4
27	21542.9	405.4	67	35549.2	544.5
28	22227.8	684.9	68	36028.4	479.3
29	22539.6	311.8	69	36409.8	381.4
30	22779.0	239.4	70	36622.6	212.7
31	23377.5	598.6	71	36898.9	276.3
32	23743.5	366.0	72	37185.0	286.1
33	24180.6	437.2	73	37485.6	300.6
34	24546.2	365.5	74	37685.6	200.0
35	24759.2	213.0	75	37924.0	238.4
36	25461.4	702.2	76	38263.6	339.7
37	25887.8	426.4	77	38479.3	215.6
38	26094.5	206.7	78	38787.6	308.3
39	26321.1	226.6	79	38991.8	204.2
40	26785.8	464.8	80	39191.8	200.0

Table 4.13: The radii of the PMMA dichromat with NA=0.83, which we fabricated and characterized its point-spread function (PSF.) (phase step = 400 nm) E_1 = 1.2712E+18, E_2 = 1.4714E+18, E_{node} = 2.7719E+15.

Zone	Radius (nm)	Zone Width (nm)	Zone	Radius (nm)	Zone Width (nm)
1	2177.9		41	27355.1	587.5
2	3187.1	1009.2	42	27555.2	200.0
3	6001.1	2814.0	43	27889.2	334.0
4	7495.1	1494.0	44	28192.2	303.0
5	8257.3	762.2	45	28714.5	522.3
6	8701.3	444.0	46	28930.6	216.1
7	9103.7	402.4	47	29246.4	315.8
8	10094.4	990.7	48	29570.2	323.8
9	11102.8	1008.4	49	29862.7	292.5
10	12042.7	939.9	50	30173.2	310.5
11	12664.6	621.9	51	30373.2	200.0
12	13702.2	1037.6	52	30839.2	466.0
13	14046.4	344.2	53	31065.0	225.8
14	14777.3	730.9	54	31704.1	639.1
15	15355.0	577.7	55	32043.0	338.9
16	15601.9	246.9	56	32729.4	686.4
17	15973.1	371.2	57	32938.0	208.7
18	16622.3	649.2	58	33164.3	226.3
19	17456.9	834.6	59	33364.3	200.0
20	18149.8	692.9	60	33564.3	200.0
21	18440.7	290.8	61	33764.3	200.0
22	18903.0	462.3	62	34052.8	288.4
23	19512.5	609.5	63	34300.5	247.7
24	19862.9	350.4	64	34657.1	356.6
25	20764.0	901.0	65	34979.0	321.9
26	21205.9	441.9	66	35259.4	280.4
27	21614.6	408.7	67	35807.4	548.0
28	22188.8	574.3	68	36029.7	222.3
29	22475.6	286.8	69	36231.2	201.6
30	22691.4	215.8	70	36581.8	350.5
31	23348.3	656.9	71	36812.9	231.1
32	23842.8	494.5	72	37203.3	390.4
33	24245.8	403.0	73	37440.2	236.8
34	24615.9	370.2	74	37645.4	205.2
35	24815.9	200.0	75	38152.4	507.0
36	25388.6	572.6	76	38352.4	200.0
37	25927.8	539.2	77	38552.4	200.0
38	26172.5	244.7	78	38752.4	200.0
39	26446.4	273.9	79	38952.4	200.0
40	26767.6	321.3	80	39153.9	201.6

Table 4.14: Table 4.13 continued.

Zone	Radius (nm)	Zone Width (nm)	Zone	Radius (nm)	Zone Width (nm)
81	39353.9	200.0	121	52051.2	200.0
82	39567.8	213.9	122	52251.2	200.0
83	39964.4	396.6	123	52451.2	200.0
84	40378.6	414.2	124	52651.2	200.0
85	40726.2	347.5	125	52851.2	200.0
86	40926.2	200.0	126	53051.2	200.0
87	41635.0	708.9	127	53251.2	200.0
88	41967.8	332.7	128	53451.2	200.0
89	42429.3	461.6	129	53651.2	200.0
90	42629.3	200.0	130	53851.2	200.0
91	42928.8	299.5	131	54051.2	200.0
92	43445.2	516.4	132	54251.2	200.0
93	43658.6	213.4	133	54451.2	200.0
94	43858.6	200.0	134	54651.2	200.0
95	44058.6	200.0	135	54851.2	200.0
96	44258.6	200.0	136	55051.2	200.0
97	44458.6	200.0	137	55251.2	200.0
98	44658.6	200.0	138	55451.2	200.0
99	44858.6	200.0	139	55651.2	200.0
100	45130.0	271.4	140	55851.2	200.0
101	45330.0	200.0	141	56051.3	200.0
102	45530.0	200.0	142	56251.3	200.0
103	45730.0	200.0	143	56451.3	200.0
104	45930.1	200.0	144	56651.3	200.0
105	46130.1	200.0	145	56851.3	200.0
106	46330.1	200.0	146	57051.3	200.0
107	46630.1	300.0	147	57251.3	200.0
108	46916.7	286.6	148	57451.3	200.0
109	47116.7	200.0	149	57651.3	200.0
110	47316.7	200.0	150	57851.3	200.0
111	47516.7	200.0	151	58051.3	200.0
112	47716.7	200.0	152	58251.3	200.0
113	47916.7	200.0	153	58451.3	200.0
114	50623.7	2707.0	154	58651.3	200.0
115	50851.1	227.4	155	58863.6	212.3
116	51051.1	200.0	156	59066.5	202.9
117	51251.1	200.0	157	59266.5	200.0
118	51451.1	200.0	158	59466.5	200.0
119	51651.1	200.0	159	59666.5	200.0
120	51851.2	200.0	160	59866.5	200.0

4.5.1 The Fabrication Process

The zones were patterned in PMMA with single-pixel lines on a double-side polished fused-silica substrate using a RAITH-150^{TM} SEBL system at an acceleration voltage of 30 keV, as illustrated in Fig. 4-5(a). A thin layer of aluminum was put on top of the PMMA to reduce sample charging during exposure. The phase shift between alternating zones was 0.997 λ_1 (399 nm for λ_1 = 405 nm) for numerical apertures NA = 0.55, 0.83, and 0.846 λ_1 (338 nm) for NA=0.7. Proximity-effect correction was done empirically by reducing zone widths or by adjusting electron-beam exposure doses. After patterning, the aluminum was removed in 0.26N tetramethylammonium hydroxide for 60 s before development of PMMA in a 1:3 methylisobutylketone/isopropanol (MIBK/IPA) solution for 90 s. Fig. 4-5(b)-4-5(d) show scanning-electron micrographs of completed dichromats.

In order to demonstrate that absorbance modulation can be easily parallelized with diffractive optics to achieve practical throughputs, arrays of dichromats were fabricated on the same glass substrate. The spacing between dichromats is chosen to be large enough that signals generated from the dichormats can be easily separated by the collection optics in the case of imaging. Figure 4-6(a) shows an array of dichromats with various numerical apertures (NAs) ranging from 0.83 to 0.22 with a layout shown in Fig. 4-6(b). Figure 4-6(c) shows a zoomed in view of the dichromats with NA=0.83 (160 zones), NA=0.7 (80 zones), NA=0.42 (20 zones), NA=0.30 (10 zones), NA=0.22 (5 zones), all in 400nm-thick PMMA. Except for the NA=0.7 dichromat in Figure 4-5, which was fabricated and characterized according to an earlier design, all other dichromats were fabricated in 400nm-thick PMMA and are shown in Fig. 4-6(a). These dichromats designs were obtained with the PMMA step height, h, removed from the vector of optimization parameters. h is set to give exact π phase shift at λ_1 to enable fabrication of different dichromat designs or conventional Fresnel zone-plates on the same substrate. Figure 4-6(d) shows an array of dichromats and zone-plates fabricated side by side for comparison. The dichromats were spaced by a distance of 140μm in the X and 160μm in the Y direction. To ensure proper gapping, a set of

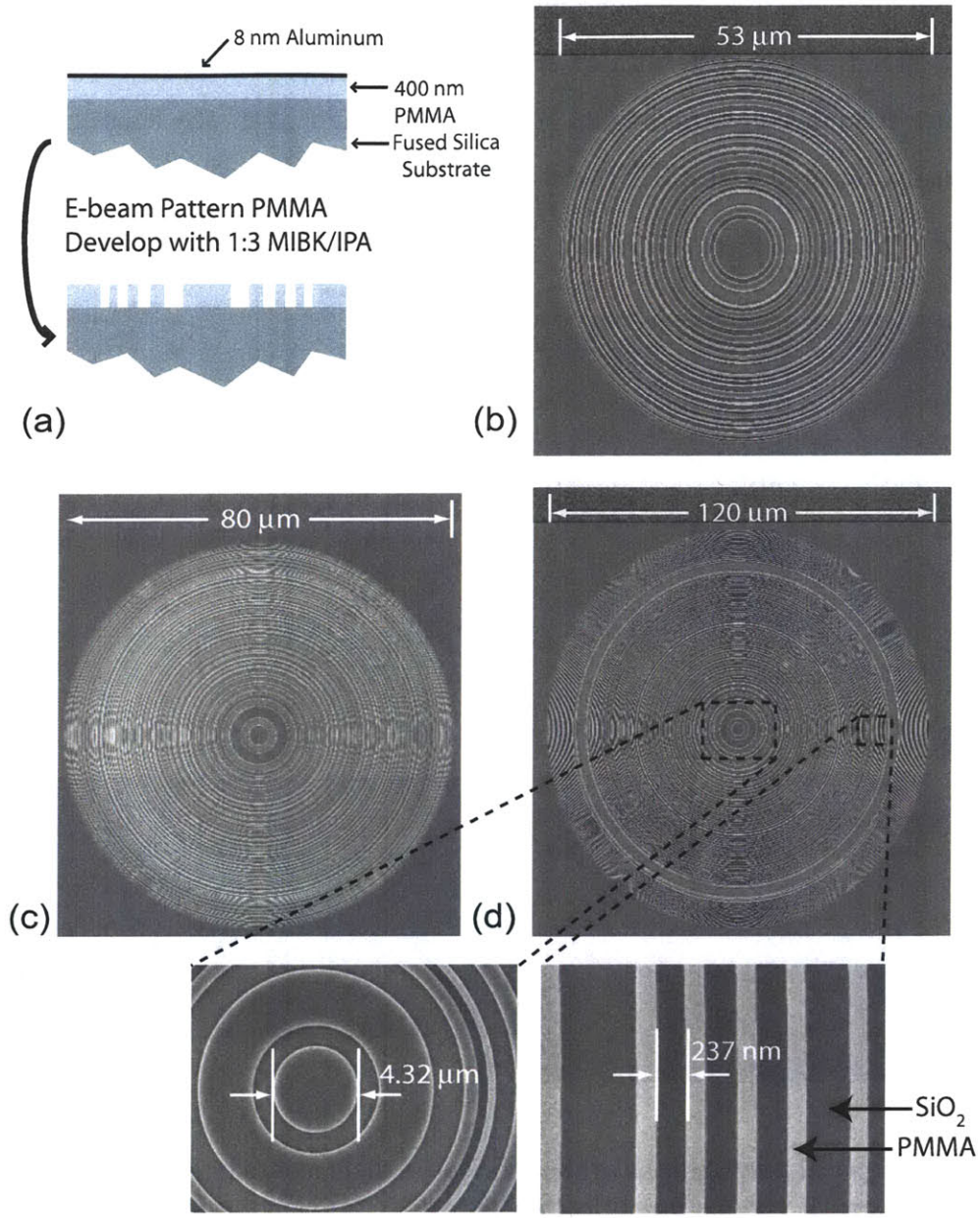

Figure 4-5: (a) Schematic of the fabrication process. PMMA was spun on top of a fused-silica substrate, and a thin layer of aluminum evaporated on top as a conduction layer. Exposed PMMA was developed in a solution of MIBK and IPA in the ratio of 1:3. Scanning-electron micrographs of the fabricated dichromats with (b) NA=0.55, (c) NA=0.7, and (d) NA=0.83. The moirè artifacts are a consequence of image formation via scanning. Insets at the bottom show magnified images of the NA=0.83 dichromat. Note that the sequence of zone radii differs significantly from that of a Fresnel zone plate.

three satellite Fresnel zone plates, designed for λ=633nm with an NA of 0.85 and the same 40μm focal length as the dichromats, was fabricated on the same glass substrate and brought into focus with a linear actuator before illuminating the dichromats for absorbance modulation experiments [123]. The 633nm light was chosen because it can be easily detected with CCD cameras and has minimal effects on most photoresist we use.

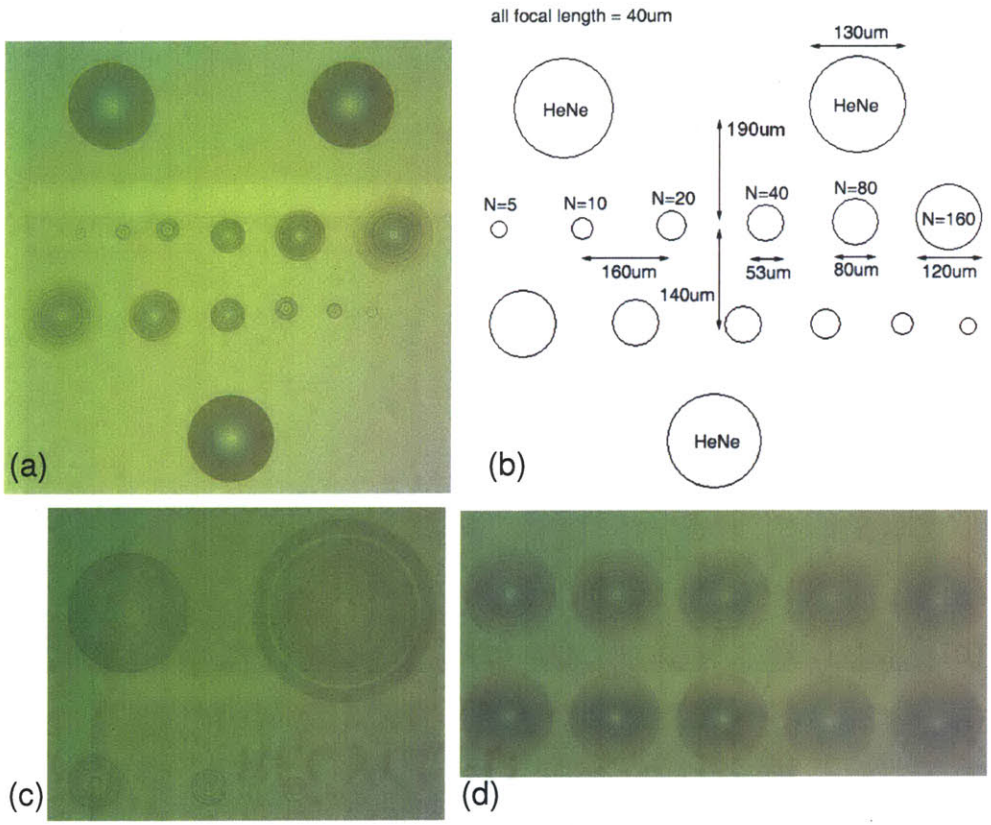

Figure 4-6: Optical micrographs of the fabricated dichromats in PMMA. (a) An array of dichromats with various numerical apertures (NAs) ranging from 0.83 to 0.22 and three satellite zone-plates around the array. (b) Layout and scale for the array in (a). (c) A zoomed in view of the dichromats with NA=0.83 (160 zones), NA=0.7 (80 zones), NA=0.42 (20 zones), NA=0.30 (10 zones), NA=0.22 (5 zones). (d) An array of dichromats and zone-plates fabricated side by side for comparison.

Finally, a self-aligned wet etch technique first proposed by T.A. Fulton and G.J.Dolan [124, 125] is applied to cover areas outside of the dichromats with chrome absorbers to reduce effects of stray lights coming from outside of the dichromat regions [126, 127]. The Fulton-Dolan process can selectively control whether etching, electroplating, or

anodizing takes place by applying a bias voltage onto the target metal substrate. More conveniently, a bias voltage can be applied simply by electrically contacting the substrate with proper metals without use of any external voltage source. In this particular case, the target metal is chrome and the biasing metal is copper. First, an 80nm-thick layer of chrome is e-beam evaporated to cover the entire dichromat substrate. Copper tape and copper clips are then brought into contact with the chrome layer outside of the dichrmat regions to passivate chrome etching at the contacted regions. When immersing the substrate into a chrome etchant, CAN-800 brushfire etchant from Transene company, Inc. (800 g/gallon ceric ammonium nitrate), for 2 minutes, only isolated chrome features over the dichromats are etched away. This process results in a chrome absorbing film with open circles right on top of the active DOEs without any alignment process. Figure 4-7 shows optical micrographs of a set of NA=0.7 dichormat array before and after the Fulton-Dolan process.

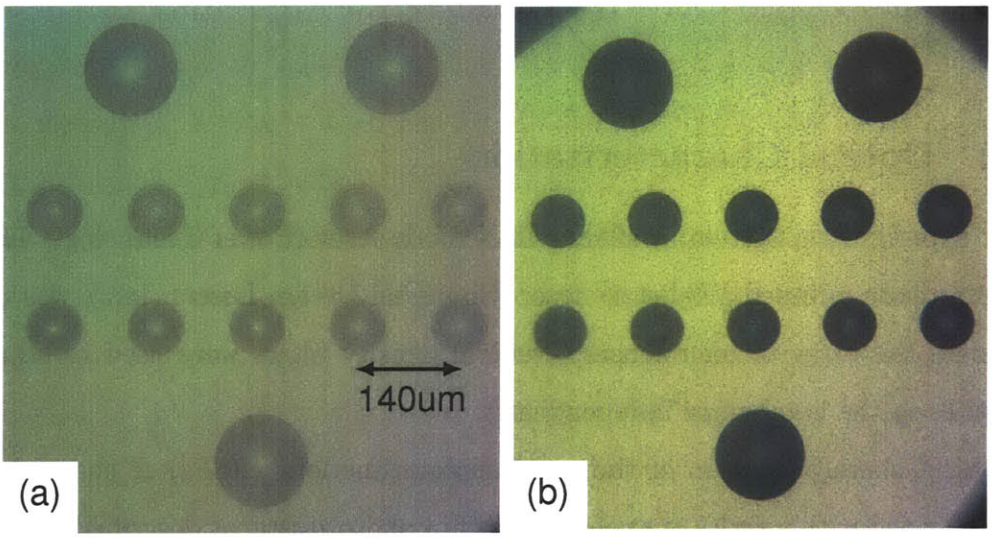

Figure 4-7: Optical micrographs (imaged in reflection) of a set of NA=0.7 dichormat array before and after the Fulton-Dolan proces.

The Fulton-Dolan process is not very sensitive to etching time, but extra care must be taken when drying the sample. Since the etchant is aqueous-based, a water rinse is needed before disconnecting copper contacts. When blow-drying water from the surface of the dichromats, the blow-drying force in conjunction with the

surface tension of water can pull PMMA zones up from the surface or pull zones together, especially those with high aspect ratios. To alleviate adhesion problems, glass substrates are RCA cleaned beforehand and UV-Ozone cleaned right before PMMA spinning. Blow-drying is done gently by avoiding direct blowing on the optics.

One last thing to note about the Fulton-Dolan process with PMMA optics is the high stress in chrome films on PMMA. In the right optical micrograph in Fig. 4-7, cracks in the chrome film around satellite Fresnel zone-plates at the top shows evidence of stress. In some extreme cases, this stress can cause the chrome film to detach locally from underlying PMMA optics and result in significant offset between the optics and the aperture in chrome, as shown in Fig. 4-8. As shown in Fig. 4-8(a), the resuilting film was already under high stress immediately after the Fulton-Dolan process. The chrome thickness used for these samples was only 50 nm, but was probably e-beam evaporated too quickly. Subsequent chrome evaporations for this process were deliberately done at a slower evaporation rate and the stress level was greatly reduced.

4.5.2 Optical Characterization

In our experimental system, a 405nm GaN diode laser (Power Technology, Inc) and a 532nm diode-pumped solid-state laser (Shanghai Dream Lasers, Inc.) was used to illuminate the array of microlenses. Neutral density filters were used to adjust the intensities in the two beams independently.

Optical characterization of the point-spread-functions (PSFs) of the dichromats was performed with a series of exposures in thin photoresists, as described in reference [9]. A threshold model of photoresist development, which assumes that below a certain dose a positive photoresist will not be developed and a negative photoresist will be removed, was assumed. By exposing the focal spots over a range of doses and measuring the corresponding radii of the spots, the PSF can be constructed.

The substrate for PSF characterization was a silicon substrate spin-coated with 200 nm of antireflection coating (BarLi Microchemicals, Germany) and 100nm of PFI-88, a positive-tone I-line photoresist from Sumitomo Chemicals. Exposure times

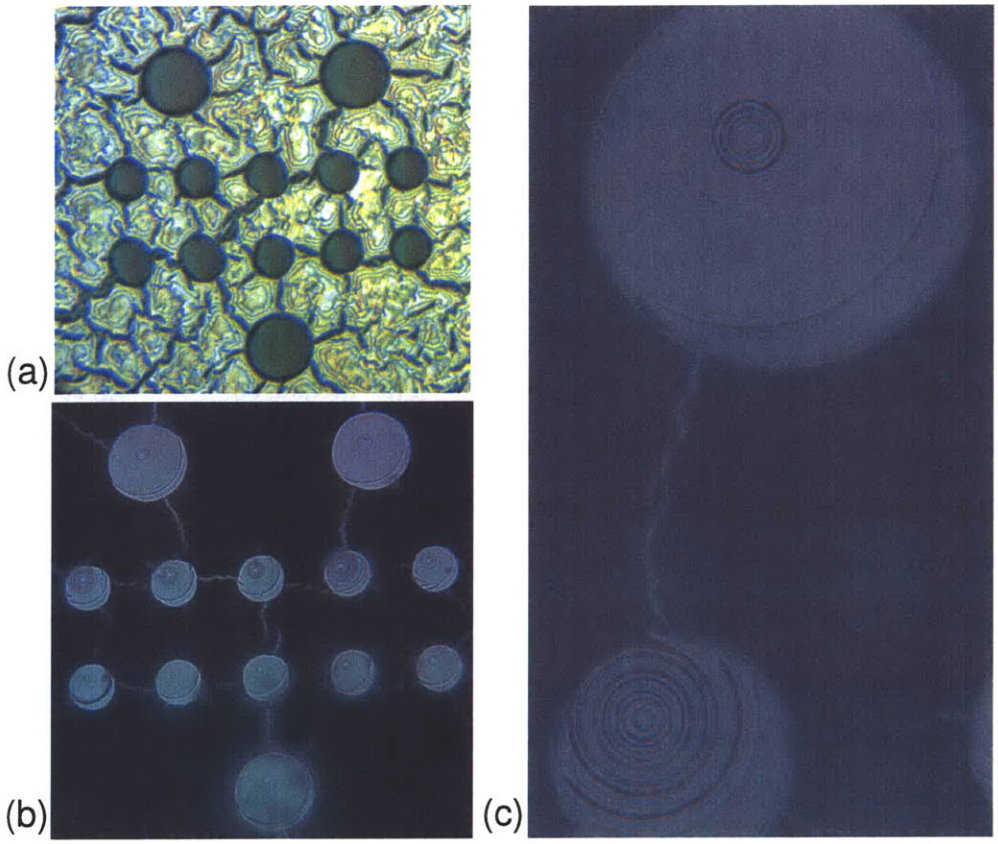

Figure 4-8: The effect of highly stressed chrome layer on PMMA after the Fulton-Dolan process. (a) A set of PMMA dichromats(top row)/ zone-plate(bottom row) with 100nm Cr coating imaged in reflection after the Fulton-Dolan process. The high density of cracks shows that the Cr film is highly stressed. (b) Another PMMA fabricated with the same process and observed after six months, imaged in transmission. While the cracks and the general shape of both the Cr film and the PMMA optics were preserved, the Cr apertures became offset with respect to the underlying optics. (c) A zoomed-in optical micrograph of this displacement. The element on the top is a 0.85 NA HeNe zone plate.

ranged from 40 seconds to 90 minute. Samples were developed for 60s in CD26 (0.26N tetramethylammonium hydroxide) solution. A series of spots were recorded in the photoresist at increasing exposure times; their radii were measured using a scanning-electron microscope and plotted against the inverse of the exposure times. The measured data were then linearly scaled using a least-square fit to the simulated data as shown in Fig. 4-9. Only dichromats with NA larger than 0.55 (40 zones) are characterized due to low efficiencies of dichromats below that NA.

As expected, dichromats focus λ_1=400 nm to round spots and λ_2=532 nm to ring-shaped spots [see insets in Fig. 4-9(c)]. The measured PSFs agree well with simulation, confirming the focusing properties of the dichromat. As the NA of the dichromat increases, from 0.55 to 0.7 to 0.83, tighter focusing at both wavelengths is observed. However, it was also found that the intensity at the center of the λ_2 focal ring increases rapidly with defocus at higher NAs. Ensuring that the photoresist surface lies within 100nm of the focal plane of the dichromats was challenging. From previous sections, we learned that DOF of dichromats are smaller than Fresnel ZP with the same NA. The DOF of He-Ne ZPs is $\frac{\lambda}{(NA)^2}$=876nm which is much larger than the DOF of the 0.83 NA dichromat.

The focusing efficiency of the dichromats was not experimentally characterized for comparison to theory, but qualitatively, the exposure time (or incidicent beam power) required for lithography or imaging is much higher for dichromats than for binary-phase zone plates with the same NA. The focusing efficiencies (in theory) for the three dichromats characterize above are summarized in Table 4.15. The following focusing efficiencies were calculated by integrating the energy under the FWHM region of the PSFs and deviding this value by the incident power density (intensity times dichromat area).

Table 4.15: Table of focusing efficiencies for the dichromats characterized above. The efficiencies are low compared to that of Fresnel Zone-Plates, which is around $\sim$ 40%

Dichromat NA	Efficiency for λ_1	Efficiency for λ_2
0.55	3.7%	1.23%
0.7	5%	4.2%
0.83	2.3%	1.6%

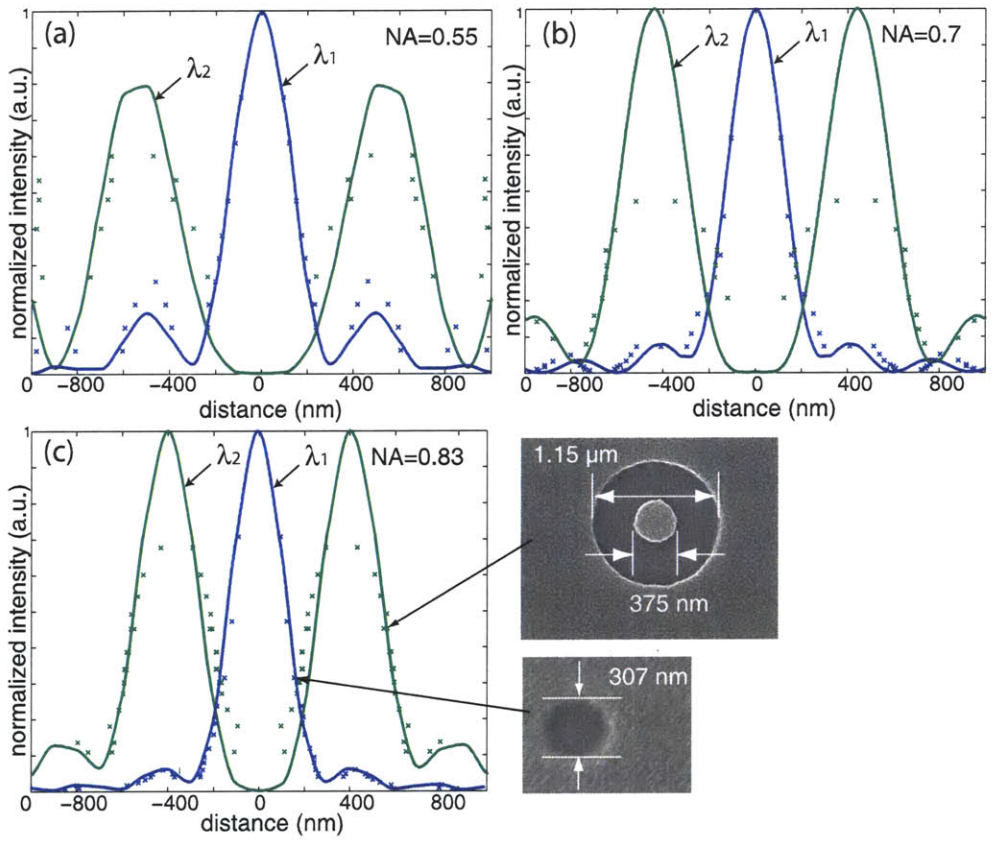

Figure 4-9: Experimental data and simulations of PSFs at the focal plane of dichromats with (a) NA =0.55, (b) NA=0.7, and (c) NA=0.83. Dots and crosses represent experimental data for λ_1=400 and λ_2=532 nm, respectively. Solid curves represent normalized theoretical PSFs. The Fresnel-Kirchhoff diffraction theory was used to calculate the theoretical PSFs. Experimental intensities were normalized using least-squares fit to simulation. Insets in (c) show typical scanning-electron micrographs of photoresist exposures for the ring-shaped intensity profile at λ_2=532 nm and a round spot λ_1=400 nm. As expected, the rings and spots are smaller at higher NAs.

4.6 Alternative Fabrication Processes

4.6.1 Different Conductive Layers

One challenge with the fabrication of phase optics on glass substrates is the charging issue during electron-beam lithography. Although a thin layer of aluminum can be coated on PMMA to prevent charging, variation in thickness or quality of this thin layer can affect the conductivity and, therefore, the focal spot size and hence clearing dose of the PMMA sample. Since zone widths of PMMA optics are sensitive to dose, we would like this conductive layer to be as reliable and repeatable as possible. In addition, evaporation and removal of the thin metal film can affect thickness of the underlying PMMA slightly, which is also undesirable. Although removal of the conductive layer is required to develop underlying PMMA, if the conductive layer is coated below the PMMA layer, removal of the conductive layer is not necessary as long as it is optically transparent. The most commonly used transparent conductive material we have access to is indium tin oxide (ITO).

The modified process flow is shown is Fig. 4-10. An 87nm-thick layer of ITO was sputtered onto glass substrates right after a 5-minute UV Ozone clean, shortly after RCA clean. A 50nm-thick layer of silicon dioxide was sputtered on top to encapsulate the ITO layer to provide insulation required for the Fulton-Dolan process. The thicknesses of the layers were designed to minimize reflection of 405 nm illumination. After spin-coating and hard baking of the PMMA layer, it was scratched to make contacts with clips. A highly conformable indium foil was placed between the clips and PMMA scratches to enable reliable contact. The sheet resistance of ITO coated samples were measured to be about 3kΩ, while that of an aluminum coated sample is usually in the MΩ range. The clearing dose for PMMA also changed from 230-240$\mu C/cm^2$ with aluminum coating to $170\mu C/cm^2$ with the ITO underlayer. Fig. 4-11 shows scanning-electron micrographs of dichromats fabricated on ITO coated substrates. The PMMA zones exhibited sharp contrast and less charging during imaging. All evidence indicates that ITO coated substrates provide better conductivity and the process affected PMMA less, resulting in lower PMMA clearing doses. Nevertheless,

the ITO process is not perfect, either. ITO is not resistant to acid cleaning, so the optics cannot be RCA cleaned even if HSQ is used as the phase shifting material.

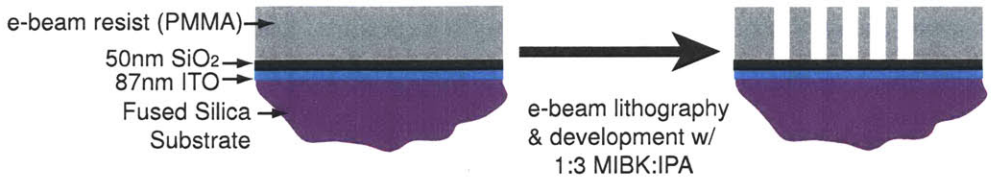

Figure 4-10: Fabrication process for diffractive optical elements in PMMA on ITO coated glass substrate. The conductive overcoating is replaced with an ITO underlayer, which does not need to be removed before or after the development of PMMA optics.

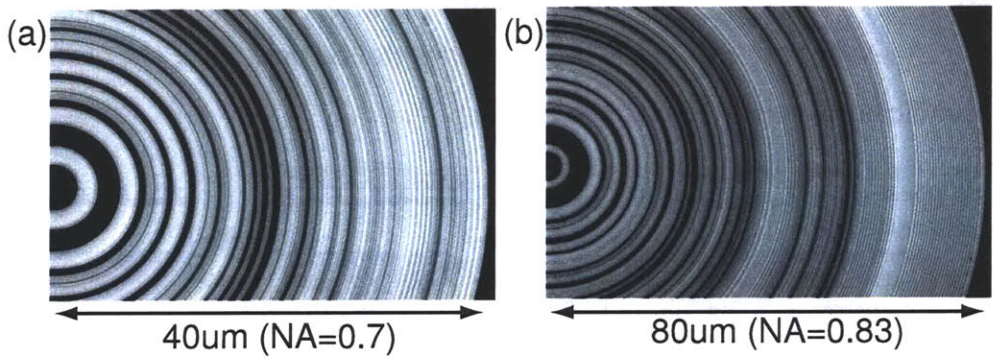

Figure 4-11: PMMA dichromats fabricated on ITO-coated glass substrates. (a) NA=0.7, (b) NA=0.83. The samples were coated with a thin conformal layer of Au-Pd for imaging. PMMA zones showed sharp contrast and less charging was observed during imaging.

Another conductive coating I investigated and used in some cases was a 60nm-thick coating of $AquaSAVE^{TM}$, a water soluble conductive polymer. The conductive polymer can be easily removed with water rinse and poses minimum effects on PMMA. The PMMA clearing dose with $AquaSAVE^{TM}$ was $200\mu C/cm^2$, but the sheet resistance of $AquaSAVE^{TM}$ is as high as 30-50MΩ or sometimes not even measurable, making it difficult to focus the electron beam on the PMMA surface.

4.6.2 Stop Layer Process

A way to avoid cleaning restrictions and minimize charging effects at the same time is to separate the lithography process from the optics-formation process. Although this process gives away the convenience of a single exposure, no pattern transfer process, the complexity it adds provides flexibility in terms of material choices and robustness for large scale fabrication.

The process flow is summarized in Fig. 4-12. The process was designed for a 450nm-thick layer of evaporated SiO_2 or hard-baked HSQ, but other materials can also be substituted. Below the phase shift layer, 133nm of MgF_2 is electron-beam evaporated on the glass substrate to serve as an etch stop and antireflection coating. On top are the lithography and transfer layers. A 50nm-thick layer of chrome is used as the hard mask for etching the underlying phase shift material as well as the conductive layer SEBL. A thin SiO_2 layer was evaporated on top as the hard mask for pattern transfer into chrome. The SiO_2 layer can be omitted if the electron-beam resist is sufficiently etch selective to chrome.

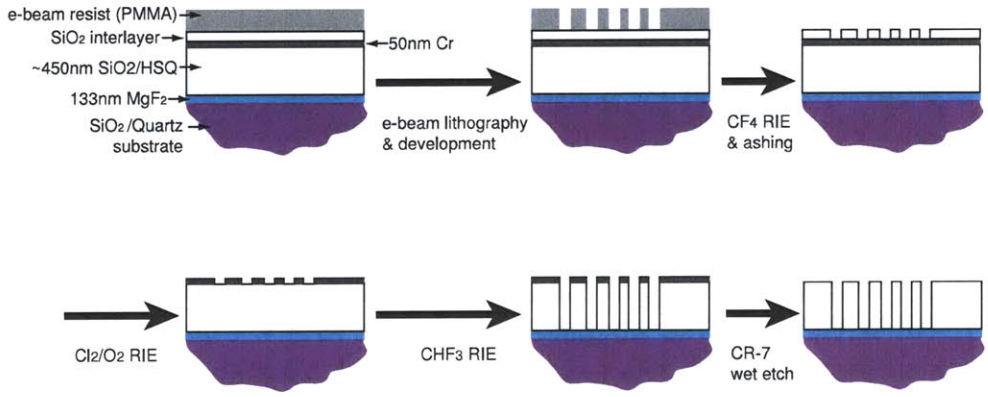

Figure 4-12: A fabrication process for diffractive optical elements with etch transfer. After SEBL patterning on the top resist layer, the pattern is transfered into a Cr hard mask for subsequent etch tranfer into a thick SiO_2 layer, which serves as the phase-shift layer. The etch stops at the MgF_2 layer. This process enables flexibility on the choice of e-beam resist and phase-shift material. Charging issues due to SEBL on glass substrates are also alleviated with the 50nn-thick Cr conductive layer.

The dichormat or zone-plate patterns are first exposed to the electron-beam resist, for example PMMA in this case. The pattern is than transferred into the SiO_2

Table 4.16: RIE conditions for the DoE etch transfer process

Etching Gas	**Tool**	**Pressure/ Flow**	**Bias**	**Power**	**Material/ Etch Rate**
CF_4	NSL Plasma Therm	10mTorr/ 15sccm	300V	100-110W	HSQ (cured)/ $\geq$ 15nn/min PMMA/35nm/min
Cl_2/O_2	SNL RIE	5mTorr/ 20sccm, 10sccm	345V	100W	Cr/ 5-6nm/min HSQ (cured)/ 6nm/min
CHF_3	NSL Plasma Therm	20mTorr/ 15sccm	200V	300-400W	HSQ (cured)/ 70-75 nm/min Cr $\leq$ 7 nm/min MgF_2/ 2.3 nm/min

interlayer with CF_4 reactive-ion etching (RIE). PMMA is not very etch resistant and etches at 35nm/min with CF_4. In order to sustain a 4-minute etch into 60nm of SiO_2, a 200nm-thick layer of PMMA was used for SEBL. After ashing excess PMMA off the SiO_2 surface, the 50nm-thick Cr layer is etched with the 60nm SiO_2 mask using Cl_2/O_2 (2:1) RIE. The last etching step into 450nm of HSQ uses CHF_3 RIE to ensure straight sidewall profiles. Finally, the remaining Cr layer is removed with the CR-7 enchant from OM Group, containing ceric ammonium nitrate and perchloric acid. The etching conditions for this process are summarized in Table 4.16. Cured HSQ was used interchangably with SiO_2 in this test. Each step was verified separately, but full integration of the entire process would require additional work.

Chapter 5

Experimental Demonstration of Absorbance Modulation with Point-wise Exposure and Image Acquisition

In the last chapter, we experimentally verify that dichromats, a new class of circular-symmetric diffractive-optical lenses, generate, in the same focal plane, focal spots for one wavelength and ring-shaped spots with central nodes for another wavelength. In the first part of this chapter, we will use these dichromats to illuminate the AML and demonstrate point-spread-function (PSF) compression of the transmitted focal spot.

5.1 Absorbance Modulation in Lithography

Following first demonstrations of absorbance modulation lithography with line exposures in Chapter 3, our goal is to extend that line spread function compression effect to point-spread-function compression to enable high resolution patterning of arbitrary patterns in a point-by-point manner [90].

5.1.1 Optical Setup

The optical setup is the same as that used for optical characterization in Chapter 4. A schematic of the optical setup is shown in Fig. 5-1. Light from a 405 nm GaN diode laser was combined with that from a 532 nm diode-pumped solid-state laser using a dichroic mirror. A separate Michelson interferometer, using a He-Ne laser, ensured that the surface of the dichromat substrate and the underlying sample were parallel. The focal spots and rings produced by the dichromats are self-aligned, so the two normally incident laser beams only need to be roughly aligned with the dichromats.

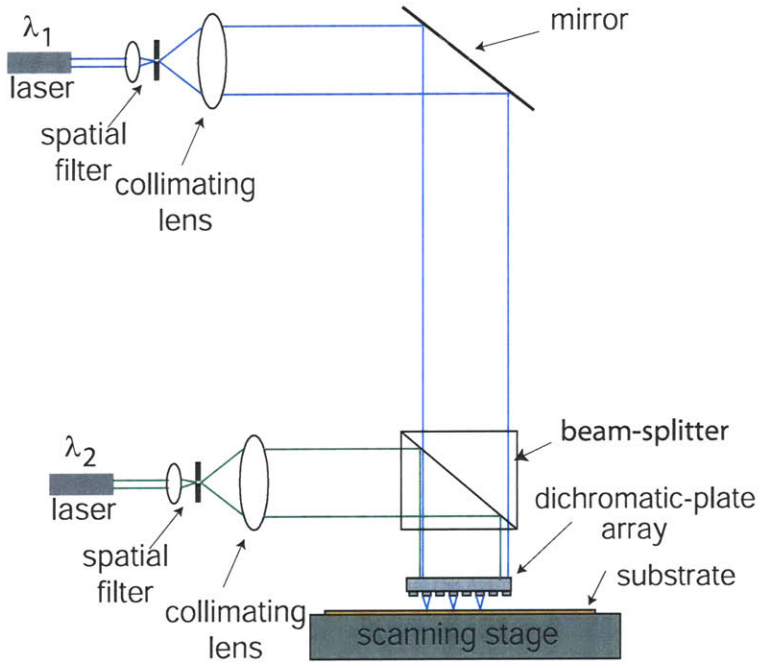

Figure 5-1: Schematic of the optical setup for absorbance modulation in lithography with dichromats that focuses the λ_1 beam into round-shaped spots and the λ_2 beam into doughnut-shaped spots.

In order to estimate the intensity ratio of the two beams illuminating the dichromats, intensity profiles of the beams at the two wavelengths were measured with a CCD camera. Since the area covered by the microlenses was much smaller than that of both the beams, we assumed that intensity incident on the microlenses was uniform. By measuring the total power within these beams, we deduced the power density incident on the dichromat array.

The definition of intensity ratio used in the first half of this chapter is slightly different from that used for simulations in Chapter 2. In Chapter 2, the input intensity ratio was defined as the ratio of the input illumination intensities before being focused down by the dichromats. In the following sections, the intensity ratio is defined as the peak intensity ratio between focused dichromat PSFs at λ_1 and λ_2. To compute the peak intensity in the focal plane of each microlens, we first simulated the normalized focal intensity distribution using a scalar Fresnel-Kirchoff diffraction formulation. Then, the peak intensity, I_{peak}, was computed as

$$I_{peak} = I_{in} \frac{\pi R^2}{\int_0^\infty I_{PSF}(r) r dr}, \tag{5.1}$$

where I_{in} is the incident (uniform) intensity, R is the dichromat radius and $I_{PSF}(r)$ is the normalized focal intensity distribution of the dichromat. Note that the peak intensities will be different for dichromats of different numerical apertures (NA) due to different focusing efficiencies. Conversion between the two definitions of intensity ratios is only a scalar multiplication, depending on the dichromat used. The conversion between total incident power (measured with the power meter) and peak intensity of the incident beam is 7.827×10^5 $1/m^2$ for the λ_1 beam and is 91.85 $1/m^2$ for the λ_2 beam. (This conversion ratio for the λ_1 beam is only valid before section 5.3 since the λ_1 laser source was changed and the incident λ_1 beam prfile was not recalibrated.) This focal-plane peak intensity ratio convention is used only when fitting experimental results to simulation, or when dichromats of different NAs are used at the same time.

5.1.2 Sample Preparation and Exposure Process

A different resist stack than that used in Chapter 4 was chosen to characterize the PSFs with absorbance modulation because both λ_1 and λ_2 expose PFI-88; high sensitivity for $\lambda_1 = 405$ nm but low sensitivity for $\lambda_2 = 532$ nm is required for absorbance modulation exposures. The sample substrates consisted of silicon wafers spin coated with 200 nm of antireflection coating, 150 nm of PS-4 (Tokyo OHKA Kogyo Co.,

Ltd., Japan), a negative-tone chemically-amplified photoresist, 23 nm of polyvinyl alcohol (PVA), and 200 nm of the azobenzene polymer (the AML) in that order. The PVA layer served as a barrier to protect the underlying photoresist stack from the AML solvent. The same development process as described in [96] was applied to the recording stack. The spot exposure times with the NA = 0.55 dichromat ranged from 4 to 60 min.

5.1.3 Lithography Results and Limitations

In Chapter 4, we pointed out that the depth-of-focus (DOF) of dichromats is small compared to that of Fresnel zone plates with the same NA. Since placing resist samples at the focal plane of the dichromats is essential for lithography, the small DOF made it challenging to control the exposure gap to the required precision with the relatively slowly varying confocal signals from the satellite He-Ne zone plates. As a result, PSF compression via absorbance modulation will only be discussed with the 0.55 NA dichromats, which have the largest depth of focus.

Figure 5-2 demonstrates PSF compression via absorbance modulation with 0.55 NA dichromats. As the intensity at λ_2 is increased relative to that at λ_1, the λ_1 illumination is focused, more tightly, as indicated in Fig. 5-2(a). The FWHM of the λ_1 PSF is decreased from 300 to 250 nm when the I_2/I_1 ratio is 20. In this set of experiment, two intensity ratios were used and the data for each ratio was least-squares fit separately. The two ratios, $I_2/I_1 = 2$ and $I_2/I_1 = 20$, best fit both sets of data while maintaining a roughly 10 times difference between the two ratios. Because the $I_2/I_1 = 2$ set was conducted with the λ_2 beam illuminating slightly off axis on the dichromats and the $I_2/I_1 = 20$ set was with the λ_2 beam centered on the dichromats while keeping the λ_1 beam position fixed, the 10 times difference between the two ratios was also an estimate. Experimentally, the peak intensities of the incident beams were measured to be 0.7164 W/m^2 at $\lambda_2 = 532$ nm and 5.912 mW/m^2 at $\lambda_1 = 400$ nm. This corresponds to peak intensities of 118.7 W/m^2 at $\lambda_2 = 532$ nm and 12.43 W/m^2 at $\lambda_1 = 405$ nm at the focal plane of the NA = 0.55 dichromat, which gives $I_2/I_1 = 9.5$, only a factor of 2 different from the fitted intensity ratio $I_2/I_1 = 20$. The

result shows that assumptions made on the illumination intensities were reasonable and self-consistent.

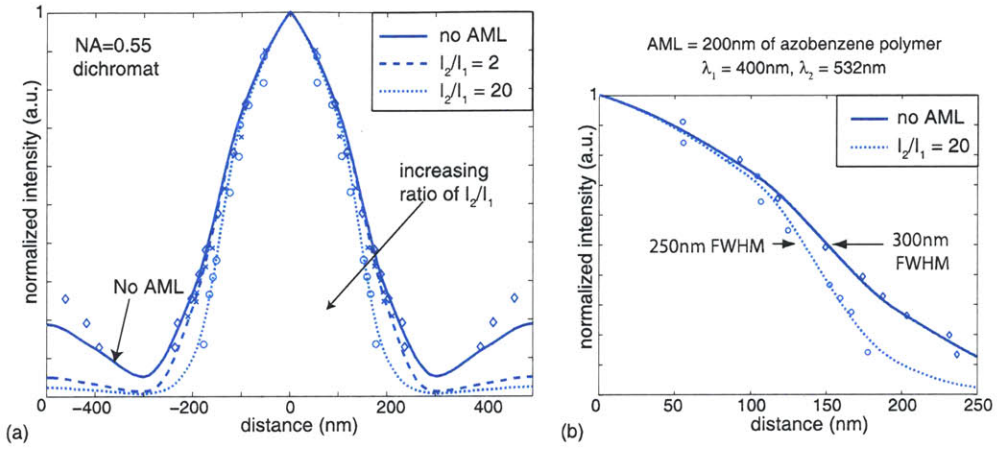

Figure 5-2: Experimental demonstration of PSF compression via absorbance modulation using a dichromat with 0.55 NA. I_1 and I_2 are the peak intensities at λ_1 and λ_2, respectively. (a) Solid curve shows the simulated PSF at λ_1 when no AML is present. The diamonds show the corresponding experimental data. The dashed curve shows the simulated PSF at λ_1 when $I_2/I_1 = 2$. The crosses show the corresponding experimental data. The dotted curve shows the simulated PSF at λ_1 when $I_2/I_1 = 20$. The circles represent the corresponding experimental data. As I_2/I_1 increases, the λ_1 illumination is focused more tightly. (b) Expansion from (a) indicating FWHM compression from 300 to 250 nm.

This result breaks the diffraction limit at $\lambda_1 = 405$ nm for a lens with an NA of 0.55, which is $\lambda_1/2NA = 362$ nm. Note that the FWHM of the PSF at $\lambda_1 = 400$ nm, before applying the absorbance modulation, is already smaller than the diffraction limit, but the PSF exhibits side lobes at more than 20% of its peak. Similar results were reported for diffractive superresolving elements[128]. With absorbance modulation, not only is the FWHM reduced but also the sidelobes. Because PS-4 is still slightly sensitive at λ_2, it was not possible to increase the ratio beyond 20 without exposing PS-4 with the λ_2 ring-shaped profile. Photoresists with higher wavelength selectivity, or AML materials and dichromat designs for a longer λ_2, should enable higher ratios and tighter focusing. Other factors, such as thermal instability of the AML, also limited the lowest feasible intensity at λ_1 and, therefore, limited the highest I_2/I_1 intensity ratio. Improving dichromat designs for larger depth of focus at high NAs would also

help achieve higher spatial resolution.

In summary, The PSF measurements matched well with simulations, verifying the design technique and the feasibility of creating optical nodes with circular-symmetric diffractive-optical lenses. We used the fabricated dichromats to experimentally demonstrate tighter focusing of light using absorbance modulation. These results will pave the way via optimized photochromic molecules to "focusing" of light beyond the diffraction barrier with applications in optical nanoscopy and nanopatterning.

5.2 Absorbance Modulation Imaging with Dichromats and Low NA Signal Collection

Absorbance modulation can be applied not only in lithography, but also for imaging. By spin-coating the AML on the imaging targets, the spatially-confined probe beam can excite a signal in the sample from a nanoscale volume. The excitation volume is not only laterally confined by absorbance modulation, but also highly confined in depth at the surface of the sample because the optical near-field decays exponentially beyond the photochromic film. Since absorbance modulation is a surface imaging technique, it may be useful for studying surface pathways, which currently use total-internal-reflection microscopy (TIRF). By stepping the sample relative to the optics and tabulating the signal as a function of position, an image similar to a conventional scanning-optical microscope is built up. The mechanism that generates the nanoscale probe, *i.e.* absorbance modulation, is distinct from the mechanism which generates the signal, *i.e.* the interaction of the probe with the sample. As a result, the two mechanisms can be independently optimized. In contrast to previous approaches, nanoscopy using absorbance modulation is not limited to fluorescence imaging. Most results presented here are based on transmitted light images. Finally, in order to achieve practical throughputs, multiple probes are necessary. Absorbance modulation is relatively easy to parallelize because high intensities are not required and only far-field optics is used. A large array of such microlenses can provide massive parallelism

and hence high throughput[90].

5.2.1 Experimental Setup

The optical system used for lithography was converted to a scanning-imaging system by adding a pair of reflective mirrors below the substrate holder and by removing the pin-chuck that had been used to hold silicon samples to allow low NA light collection, as shown in Fig. 5-3. The same 405 nm GaN diode laser (Power Technology, Inc) and 532 nm diode-pumped solid-state laser (Shanghai Dream Lasers, Inc.) were used to illuminate the dochromat array. In the first set of experiments, we used an array 10 dichromats with NA ranging from 0.83 to 0.22, as shown in Fig. 5-3 and Fig. 4-6(b). The dichromats were spaced at a distance of 140 μm in the X and 160 μm in the Y direction so that signals from different dichromats could be separated on the CCD plane even with low NA collection optics. Hence, signals were collected in parallel similar to a system described earlier for zone-plate-array imaging[90]. The sample, which consisted of features on a glass slide, was affixed (with carbon tape) on top of a piezoelectric scanning stage (Physik Instrumente). The stage had an open aperture through which the light at λ_1 scattered in the forward direction from the sample was collected by a CCD camera (Sony). Scattered light at λ_2 was filtered out. Before imaging, the surface of the dichromat array was made parallel to the sample and brought into gap by maximizing confocal signals from the satellite HeNe zone plates using piezoelectric actuators. Unlike lithography, in which gapping quality is difficult to assess before taking the sample out and developing it, imaging results directly indicate whether the illumination optics are in focus, so the gap can be adjusted for best image quality. At each scan location, the sample was illuminated simultaneously by both wavelengths, and the position and signal data were collected. The signal was acquired after a delay of 1s at each location in order to ensure that the photostationary state of the AML was reached. The sample was stepped in a raster fashion. In this implementation, the 1-s delay during signal acquisition limits the overall imaging speed. This can be speed up by increasing the incident light intensities or using an AML material with faster settling times. Although these preliminary experiments

were slow, they demonstrate the feasibility of parallelization and the potential for faster imaging.

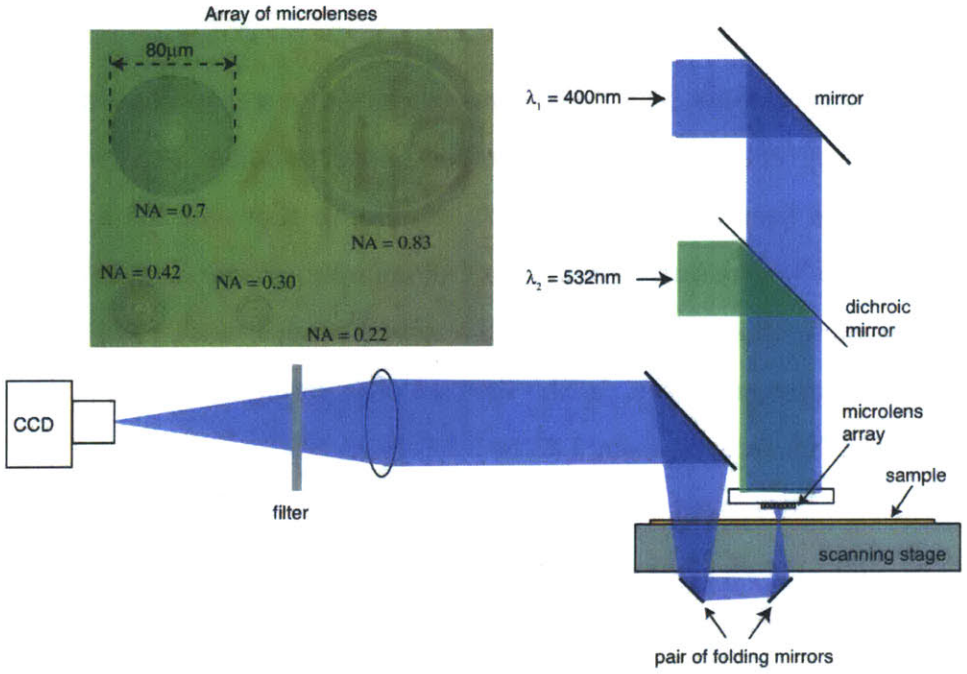

Figure 5-3: Schematic of the absorbance-modulation imaging (AMI) system. The inset on the left shows an optical image of the microlenses of various numerical apertures (NAs). The scanning stage has a clear aperture enabling the transmitted light to be collected. The microlenses are spaced such that the signals are read out in parallel. The λ_2 beam is not shown past the sample for clarity. In practice it is blocked by a filter from entering the CCD.

5.2.2 Imaging Results with Resolution Standards

The first type of resolution standard we imaged contained dense line/space patterns. Gratings of period 400 nm and 500 nm were patterned on glass slides using interference lithography. A tri-layer resist stack was used for interference lithography consisting of a 100 nm layer of PFI-88 (Sumitomo, positive-tone i-line resist) with 30 nm of silicon dioxide interlayer on 200 nm thick antireflection coating (ARC). After creating grating patterns in the ARC layer, the pattern was transferred into a 50-nm-thick film of chromium using a lift-off process with the EKC-265 striper at 85°C. The sample was spin-coated with a 220 nm thick film of the azobenzene AML for absorbance

modulation imaging. Figure 5-4(a) shows the images of a 500-nm-period line/space pattern obtained using a dichromat of NA = 0.83. On top is the image taken with only λ_1 (λ_1 = 405nm) illumination. The bottom image is of the same region but taken with both λ_1 and λ_2 (λ_2 = 532 nm) illumination. Even at relatively low peak intensities at λ_2 (0.4 W/m^2), it is clear that the image contrast is improved. Figure 5-4(b) shows the averaged linescan of the images in Fig. 5-4(a). The solid line shows the scan with only λ_1 illumination, whereas the dashed line depicts the scan with both λ_1 and λ_2 beams. Absorbance modulation clearly increases the image contrast. The contrast, K is defined as

$$K = \frac{< I_{max} > - < I_{min} >}{< I_{max} > + < I_{min} >}, \tag{5.2}$$

where $< I_{max} >$ and $< I_{min} >$ are the average maximum and minimum signal values in the peaks and valleys, respectively. Note that these images were taken in transmission, so the images are reversed in tone with respect to the object. A step-size of 25 nm was used for Fig. 5-4.

This experiment demonstrated contrast enhancement with absorbance modulation, but also revealed one drawback of this imaging setup. As seen in Fig. 5-4(a), the imaging data after applying absorbance modulation is quite noisy. The contrast is enhanced at an expense of lower signal level and therefore higher signal to noise ratio. When imaging samples with even smaller gratings periods, signal level decreases even before applying the AML due to loss of high spatial frequency components of the transmitted light. Because metal on glass samples do not scatter incident lights very strongly, signals from such samples are in general low. This not only limits the image quality, but also the level of absorbance modulation that can be applied, *i.e.* how intense the λ_2 illumination can be. One way to get around this limitation is to image samples that scatters the incident λ_1 light more strongly to the far field.

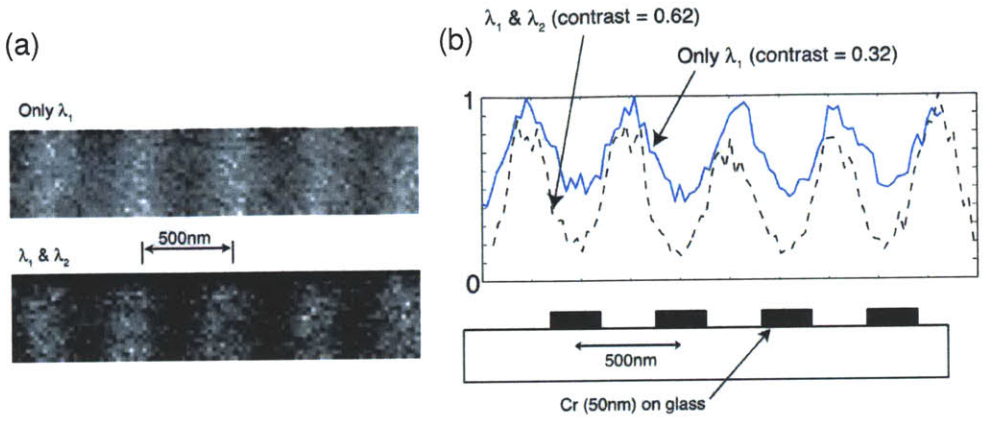

Figure 5-4: Absorbance-modulation imaging of periodic metal lines. (a) Transmitted-light images of a 500nm-period grating in a 50nm-thick layer of chromium on glass, taken with only $\lambda_1 = 405$ nm (top) and with both $\lambda_1 = 405$ nm and $\lambda_2 = 532$ nm (bottom). (b) Average linescans through the images in (a). The image contrast is increased by a factor of 2 due to absorbance modulation. The sample was over-coated with 220 nm of the azobenzene polymer. The numerical aperture (NA) of the microlens was 0.83. The signals were collected in transmission. The peak focal intensities at λ_1 and λ_2 incident on the sample were 34.6 W/m^2 and 0.4W/m^2, respectively.

5.2.3 Imaging Results with Gold Nano-particles

An alternative resolution standard was created by dispersing gold nanoparticles (British Biocell International) on a glass slide. Gold nanoparticles of various nominal sizes in water (with sodium citrate as a buffer) were dispersed either by a puddle-drop method or by spincasting onto a clean glass slide. The slide was dried in an oven at 92 degrees Celsius for about 20 minutes. Subsequently, the slides were spin-coated with a 220-nm-thick film of the azobenzene polymer.

Figure 5-5 shows images of nanoparticles on glass when the intensity ratio of λ_2 to that at λ_1 is large enough to overcome the far-field diffraction barrier. Figure 5-5(a) is a schematic of the sample being imaged. The sample consisted of 100-nm gold nano-particles overcoated with 220 nm of the azobenzene polymer. The images were taken with a dichromat of NA = 0.7 with only λ_1 in Fig. 5-5(b), and with both λ_1 and λ_2 in Fig. 5-5(c). Gaps between the nanoparticle clusters as small as 78 nm are resolved when both wavelengths are used. A step-size of 25 nm was used for

Fig. 5-5 and a bilinear interpolation algorithm was used to smooth the data with a pixel size of 5nm. This set of experiment showed strong contrast enhancement of the resulting image when absorbance modulation is turned on. Although the sample was prepared in a relatively random manner, making it difficult to directly compare absorbance modulation images with other high resolution images, such as scanning-electron micrographs (SEM), at the same locations, a conventional optical micrograph and an SEM image are shown in Fig. 5-6 to provide reference of the sample appearance.

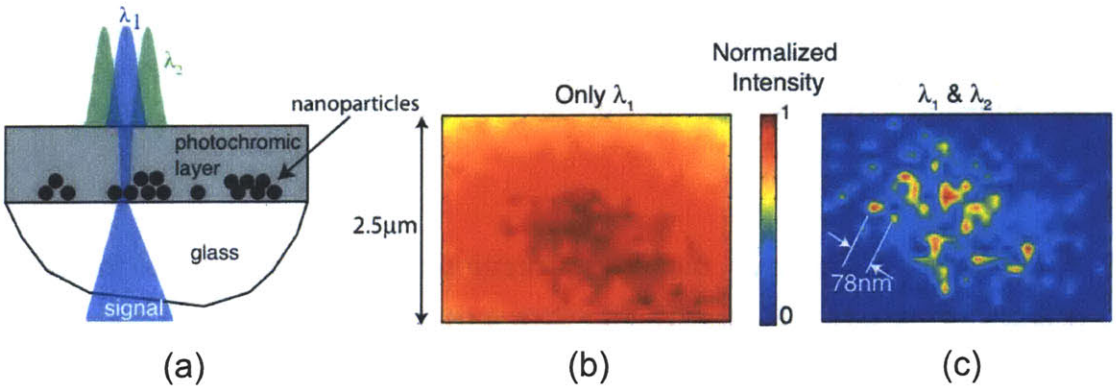

Figure 5-5: (a) Schematic of sample, a glass slide with gold nano-particles randomly dispersed on the surface. The slide was coated with 220 nm of the azobenzene polymer. Note that the signal is composed of light transmitted through the gaps between the nano-particles. (b)-(c) Images of 100 nm gold nano-particles taken by a microlens of NA = 0.7 with only $\lambda_1 = 405$ nm (b) and with both $\lambda_1 = 405$ nm and $\lambda_2 = 532$ nm (c). The peak focal intensities incident on the sample were 21.23 W/m^2 at λ_1 in (b), 13.36W/m^2 at λ_1, and 373 W/m^2 in (c), respectively.

To show how the image evolves with increasing intensity ratios, Fig. 5-7(a)-Fig. 5-7(c) show images of the same region on a glass slide dispersed with 10-nm gold nano-particles taken at several I_{2peak}/I_{1peak} ratios. As the ratio of intensity at λ_2 to that at λ_1 increases, the size of the probe decreases and finer structural details are revealed, such as a 50-nm gap between nano-particle clusters. The peak focal intensities at λ_1 and λ_2 incident on the sample were 2.95 W/m^2 and 50 W/m^2 in (a), 2.16 W/m^2 and 50 W/m^2 in (b), and 1.42 W/m^2 and 50 W/m^2 in (c), respectively. The signal levels in each image are normalized to its maximum signal level. As the

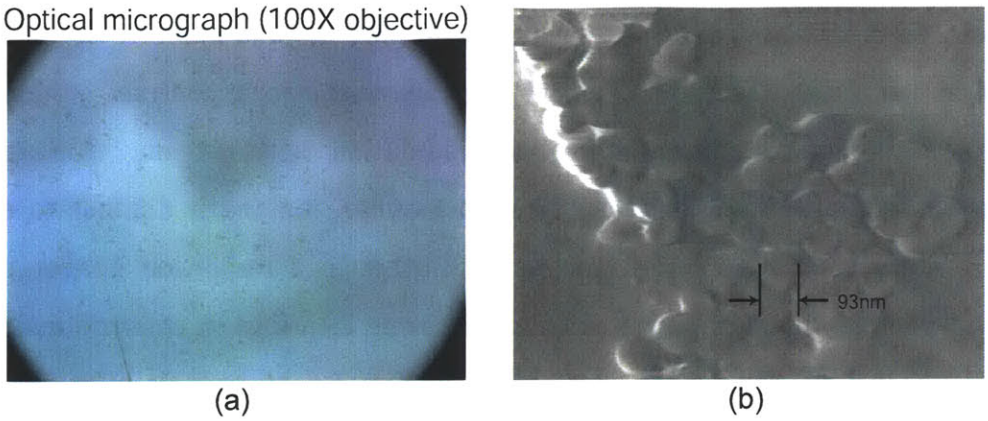

Figure 5-6: (a) Conventional optical micrograph of a glass slide dispersed with 100-nm gold nano-particles imaged with an NA = 0.9 objective lens. The height of the field is about 150 μm. Clusters of the 100-nm gold nano-particles appear as dark dots in transmission. (b) Scanning-electron micrograph (SEM) of a gold nano-particle cluster.

I_2/I_1 intensity ratio increases, nanoscale structural details are revealed at the expense of decreasing signal level and shrinking area with detectable signal, as discussed in the previous section. Figure 5-7(d) shows a magnified view of the region denoted by a white square in Fig. 5-7(c). Clusters of nano-particles are evident and some of these are denoted by black dashed lines in Fig. 5-7(d) for illustration. Notably, two (possibly) single nanoparticles spaced by a distance of 40 nm is clearly resolved. Figure 5-7(e) shows the intensities along the white lines in Fig. 5-7(a)-Fig. 5-7(c). A step-size of 25 nm was used for Fig. 5-7 and a bicubic interpolation from the raw image data was used in Fig. 5-7(d) and (e).

Finally, we show that multiple images can be collected simultaneously with three different dichromats, thereby demonstrating parallelism. Figure 5-8(a) shows a schematic of the parallel-imaging system and a typical signal array pattern on the camera. A Labview control program was constructed with the IMAQ library to synchronize stage scanning and data acquisition from the CCD. Multiple regions of interest (ROIs) can be defined before starting each scan by drawing circular regions on a CCD image at start. The average intensity value at each frame within the ROIs are then saved and plotted into a 2D-matrix with the corresponding stage positions. Figures 5-8(b)-(d)

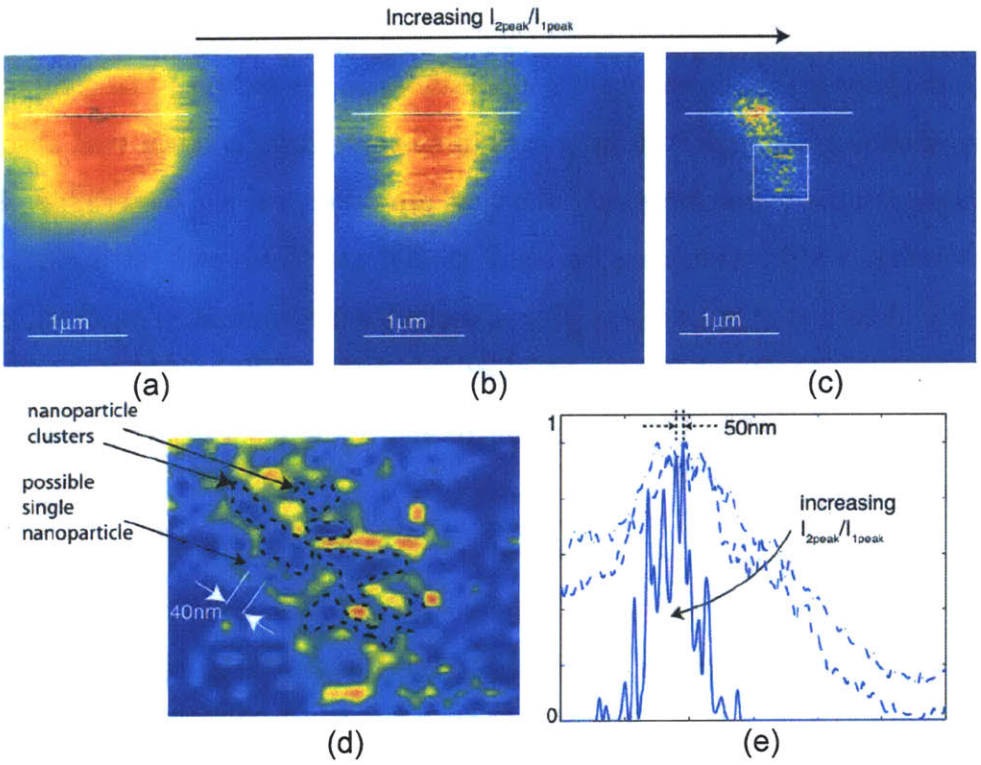

Figure 5-7: (a)-(c) Images of 10-nm gold nano-particles dispersed on a glass slide taken with a NA = 0.55 dichromat. The peak focal intensities at λ_1 and λ_2 incident on the sample were 2.95 W/m^2 and 50 W/m^2 in (a), 2.16 W/m^2 and 50 W/m^2 in (b), and 1.42 W/m^2 and 50 W/m^2 in (c), respectively. As the ratio of the intensity at λ_2 to that at λ_1 increases, nanoscale structural details are revealed. (d) Magnified image of the area within the white square in (c). Some possible geometries of nanoparticle clusters are outlined with black dashed lines. (e) Signal cross-sections through white lines shown in (a)-(c). When the intensity ratio of λ_2 to λ_1 is sufficiently high, structures that are 50 nm apart are revealed.

show images of three different regions of one gold nano-particle sample with 10nm gold nano-particles. The images in the top row were taken with only $\lambda_1 = 405$nm and the images in the bottom row depict the corresponding regions that were imaged with both λ_1 and λ_2 illuminations. The images in (b) and (c) were obtained with two different dichromats, each of NA = 0.83, while those in (d) were obtained with a third microlens of NA = 0.55. Note that the signal represents light that transmits through the gaps between clusters of nanoparticles on the sample. The comparison of the corresponding top and bottom images in all three cases confirm that absorbance modulation is able to resolve structures that are otherwise not discernible. For example, in Fig. 5-8(d), structures as small as 200 nm (full-width at halfmaximum) are clearly identified only after applying absorbance modulation. The signal levels in each image are normalized. The peak focal intensities at λ_1 and λ_2 incident on the sample were 34.6 W/m^2 and 334 W/m^2 in (b) and (c), and 3.4 W/m^2 and 37.2 W/m^2 in (d), respectively.

5.3 Absorbance Modulation Imaging with Dichromats and High NA Signal Collection

In the pervious section, we showed that by imaging gold nano-particles that interact more strongly with the $\lambda_1 = 405$ nm incident illumination, the signal strength was enhanced and more details of the signal pattern were resolved. Nevertheless, gold nano-particle samples are rather random, so cannot really act as resolution standards that provide definite sizes of reference. In this section, we explore other types of resolution standards as imaging targets for absorbance modulation with enhanced signal level by improving signal collection efficiency. Since the solid angle of signal collection is proportional to NA square, increasing signal collection NA significantly improves collection efficiency. A schematic of the modified optical setup is shown in Fig. 5-9. The original supportive stage below the piezo-electric scanning stage was replaced with a linear translation stage with an open aperture. Linear translation

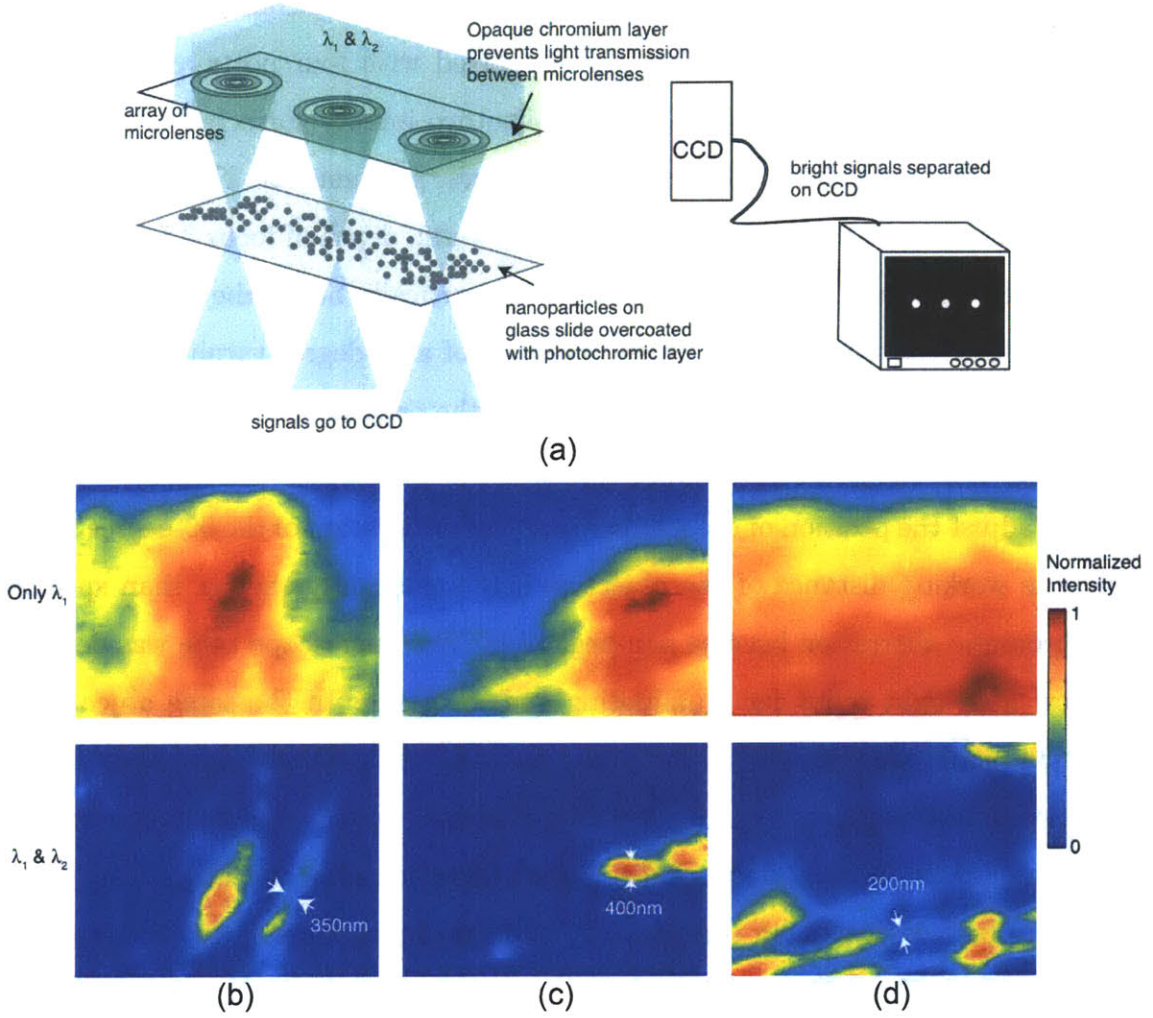

Figure 5-8: (a) Schematic of parallel imaging. The array of microlenses is illuminated by both wavelengths. Each microlens forms a focused node-spot pair that probes the sample. The transmitted light at λ_1 (the signal) is collected on a CCD. Since the focii of the microlenses are far apart, the signals are spatially separated on the CCD camera as indicated on the schematic on the right. (b)-(d) Transmission images of 10-nm gold nano-particles randomly dispersed on a glass slide. Top row shows images taken with only $\lambda_1 = 405$ nm. Bottom row shows images of the corresponding regions on the sample taken with both $\lambda_1 = 405$ nm and $\lambda_2 = 532$ nm. The images in each column were acquired by different dichromat apertures in parallel. The images were taken with dichromats of NA = 0.83 in (b) and (c), and NA = 0.55 in (d). There was a slight shift in the sample between the scans shown in the top and bottom rows. The step-size for the top row during image acquisition was 100 nm and was 25 nm for the bottom row. A bilinear interpolation algorithm was used to smooth the data with a pixel size of 5 nm.

was required for both low NA and high NA setups because the travel range of the piezo-electric scanning stage is only 200 μm. In order to setup collection optics below the scanning stage, the linear stage was supported with four posts at each corner and the illumination optics was raised. A Mitutoyo infinity-corrected long working distance microscope objective lens (Edmund Optics, part number: NT46-144) with a working distance of 33.5 mm, resolving power of 1 μm, depth of focus 3.5 μm, field of view of 0.88-0.12 mm, and NA of 0.28 is set up vertically below the scanning stage. The objective lens was mounted on an open end of a 90-degree turning mirror cube and the other end of the cube was screwed into the original 7x zoom tubing for the microscope objective. The tube is then mounted from the side on a XYZ translational stage to adjust the position of the microscope. Due to addition of the 90-degree mirror turn, the working distance of the objective lens was slightly shorter than specified and chromatic aberration became non-negligible. The microscope tube was designed to allow fiber light input for illuminating and imaging with the same objective in reflection mode. This was useful for coarse focusing of the microscope objective and for locating specific regions on the sample for imaging prior to transmission imaging with the scanning λ_1 = 405 nm probe. After coarse focusing and alignment, the λ_1 = 405 nm laser is used align the dichromats, the imaging targets, and the Mitutoyo objective lens. Flood illumination of the λ_1 = 405 nm light on the imaging target was used to focus the Mitutoyo objective lens.

In addition to modifying the collection optics for higher NA collection, the camera was upgraded to a Clara interline CCD camera from Andor Technology (Sony ICX285 chip, 1360 x 1024 pixel, 6.45 μm pixel pitch, fan cooled, maximum read out rate = 11.6 frames/sec) for higher sensitivity and lower noise. For all images in the following sections, the camera was operated at -10 °C with the fan off to avoid vibration issues. The camera also supports options to bin neighboring pixels for faster data readout and frame rates. Since only the average intensities within the ROIs are of concern, binning pixels on the CCD camera enhances throughput with little effect on the image quality. Pixel binning of 2 in both X and Y are used for all following experiments. Because the settling and data acquisition time of the Andor CCD camera is much

longer (0.1 s to 1 s) than the Sony CCD camera, no additional delay was applied before signal acquisition. The Labview codes that controls signal acquisition from the Andor CCD camera are documented in Appendix C.

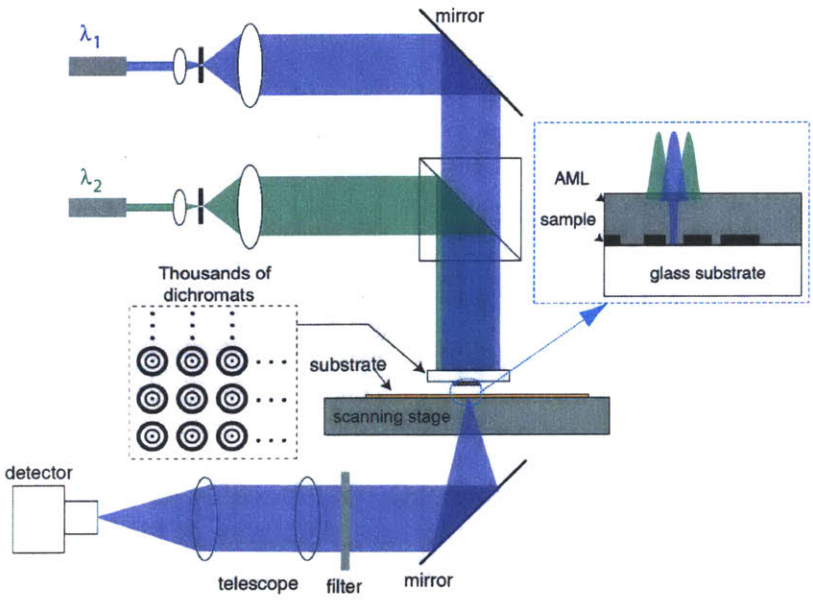

Figure 5-9: Schematic of an AMI microscope with higher NA collection optics. Each dichromat in the array focuses λ_1 to a round spot and λ_2 to a doughnut shaped spot, creating a local subwavelength aperture for λ_1 in the AML through which the underlying object is illuminated and the scattered light collected. The inset shows the schematic of a resolution test structure consisting of metal lines on a glass substrate.

5.4 Electron-Beam Patterned Resolution Standards

Two types of resolution standards, as shown in Fig. 5-10, were fabricated with scanning-electron-beam lithography on indium-tin-oxide (ITO) coated glass substrates. To reduce loss of signal during propagating in the glass substrate, an optical flat with 2-inch diameter was used as the substrate. A 60-nm layer of ITO was sputtered on top before coating 200 nm of PMMA electron-beam resist. After patterning, the PMMA was developed in a 1:3 methylisobutylketone/isopropanol (MIBK/IPA) solution for 90 s. A 50-nm layer of chrome was electron-beam evaporated on the sample and a

liftoff process conducted with N-Methyl-2-pyrrolidone (NMP) at 90 °C for 20 minutes with ultrasonic agitation at the end. This section will focus on imaging results with the resolution standard in Fig. 5-10(a).

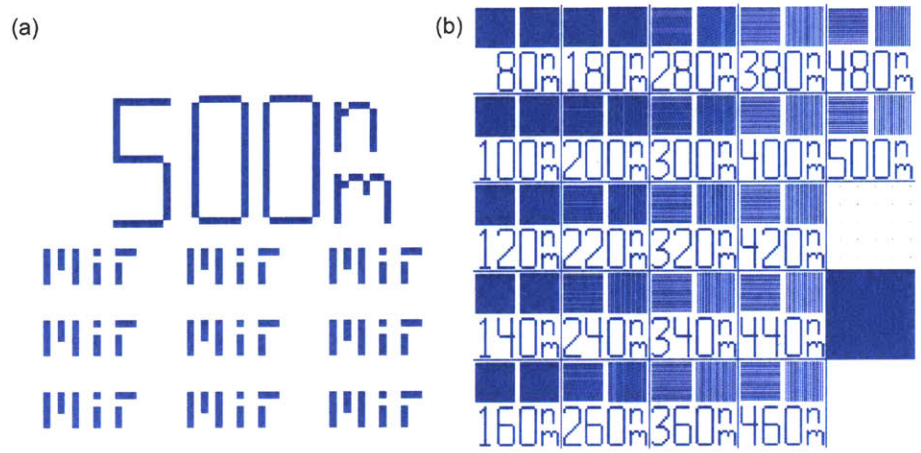

Figure 5-10: Two types of resolution standards fabricated on ITO coated glass slides with scanning-electron-beam lithography. (a) An array of MIT logos (only showing a small part of the whole pattern) with 500 nm, 200 nm, and 100 nm linewidths. (b) An array of horizontal and vertical gratings with periods from 80 nm to 500 nm. The two boxes on the lower right corner were designed be an array of small dots and array of small holes, respectively, but they did not turn out well due to very different proximity-effect levels compared to the rest of the pattern.

Figure 5-11(a) shows an MIT logo with 500-nm linewidths imaged by a scanning-electron microscope. No overcoating was required to prevent charging effect because the top surface of the sample consists of chrome patterns on an ITO layer. After imaging with the SEM, the sample was first imaged with an NA = 0.83 Fresnel zone plate, shown in Fig. 4-6(d). The Fresnel-zone-plate image, as shown in Fig. 5-11(b) showed significant rounding of the sharp chrome edges, with a radius of curvature of about half the linewidth (~250 nm). This level of rounding can be explained if we estimate the resolution for this system with $\lambda_1/2\text{NA} = 244$ nm. The sample was then overcoated with a 220 nm layer of azobenzene AML. Figure 5-11(c) shows the same MIT pattern imaged by an NA = 0.83 dichromat with the AML, but only $\lambda_1 = 405$ nm illumination. The chrome lines appears thinner when imaged with the dichromat, through the AML, indicating that the combination of dichromat plus AML PSF at

λ_1 is narrower than the PSF of the NA = 0.83 zone plate. Nevertheless, the image contrast dropped and line-edge rounding was still prominent. When imaging with both λ_1 = 405 nm and λ_2 = 532 nm illuminations, Fig. 5-11(d) showed sharper edges and less rounding at corners. This visual observation can be supported by taking two-dimensional Fourier transforms of the absorbance modulation images in Fig. 5-11. Figure 5-12 shows that Fourier transform of the NA = 0.83 dichromat image, obtained with simultaneous λ_1 = 405 nm and λ_2 = 532 nm illumination, contains a wider spread of frequency components compared to that of the NA = 0.83 Fresnel-zone-plate image. Figure 5-13 further quantifies the amplitude distribution in Fig. 5-12 by plotting the Fourier transform amplitude distributions with respect to radial frequency. The fact that the absorbance modulation image contains more high-spatial-frequency components supports the argument that the corners of the lines appears sharper with absorbance modulation. The step-size of Fig. 5-11(b)-(d) was 100 nm and no interpolation was applied to the raw data. The I_2/I_1 power ratio was 25 upon incidence.

With the well-defined resolution standard in Fig. 5-10(a), we studied the effect of defocusing in absorbance modulation by not only looking at the image contrast, but also the legibility of the MIT logo. Figure 5-14 shows a series of images taken at several gaps. The "gap" here refers to where the reflection signals of the satellite HeNe zone plates are in focus. The images show that gapping is critical to the quality of the image and that the optimal gap differs from the focal distances of the satellite zone plates. When applying both λ_1 = 405 nm and λ_2 = 532 nm illuminations, improvement on image quality is limited, which implies that the compressed FWHM was not much smaller than the linewidth, 200 nm. The step-sizes of Fig. 5-14 were 100 nm for the images at various gaps and 50 nm for the image with both λ_1 and λ_2 illuminations. No interpolation was applied to the raw data. The I_2/I_1 power ratio was 25 upon incidence.

Although it can be expected from images of the 200-nm-linewidth MIT logo that 100-nm-linewidth MIT logos likely will not be resolved, two images of a 100-nm linewidth MIT logo are shown in Fig. 5-15 for reference. The letter 'M' was illeg-

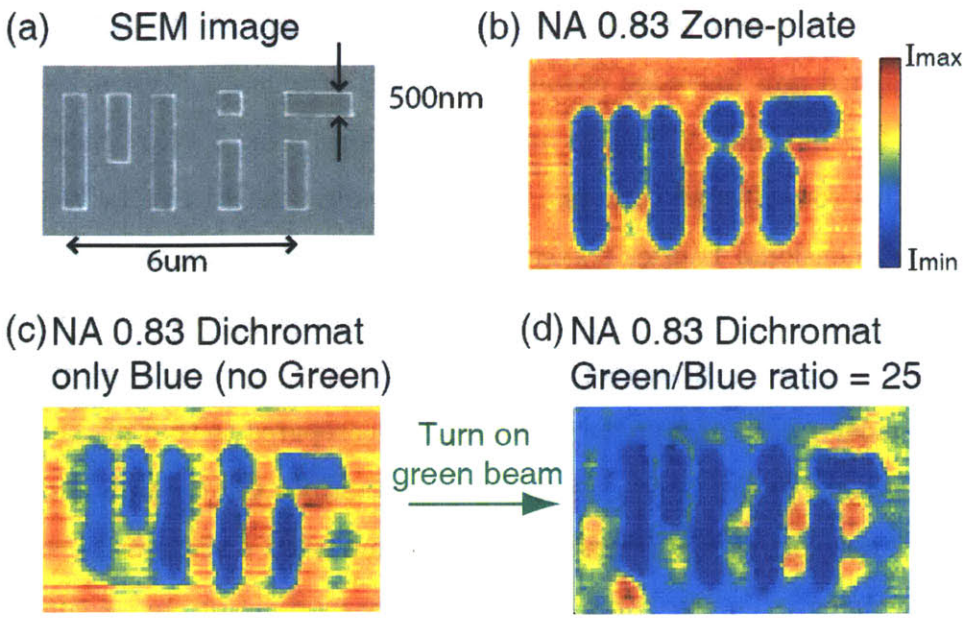

Figure 5-11: A 500-nm-linewidth MIT logo, imaged with (a) a scanning-electron microscope (b) an NA = 0.83 Fresnel zone plate (c) an NA = 0.83 dichromat with only $\lambda_1 = 405$ nm illumination (d) an NA = 0.83 dichromat with both $\lambda_1 = 405$ nm and $\lambda_2 = 532$ nm illumination.

ible and the letter 'I' became unclear. The step-size was 50 nm for Fig. 5-15 and no interpolation was applied to the raw data. The I_2/I_1 power ratio was 25 upon incidence.

A Note on PSF Characterization and Node Quality at λ_2

One possible reason for limited PSF compression with absorbance modulation is non-ideal λ_1 and λ_2 incident PSF profiles generated by the dichromats. As discussed in Chapter 2, PSF scaling with intensity ratios is very sensitive to nonzero illuminations at the central node of the λ_2 profile. Although the PSF was characterized with photoresist exposures in Chapter 4, photoresist exposures are only best for mapping high slope portions of the PSF, for example, the waist of the center lobe of the PSF. Photoresist characterization at the peaks or nodes of the PSF is difficult due to small process latitudes, *i.e.* the exposure time needs to be adjusted such that the clearing dose equals the dose at the peaks or the nodes. The resulting feature sizes would

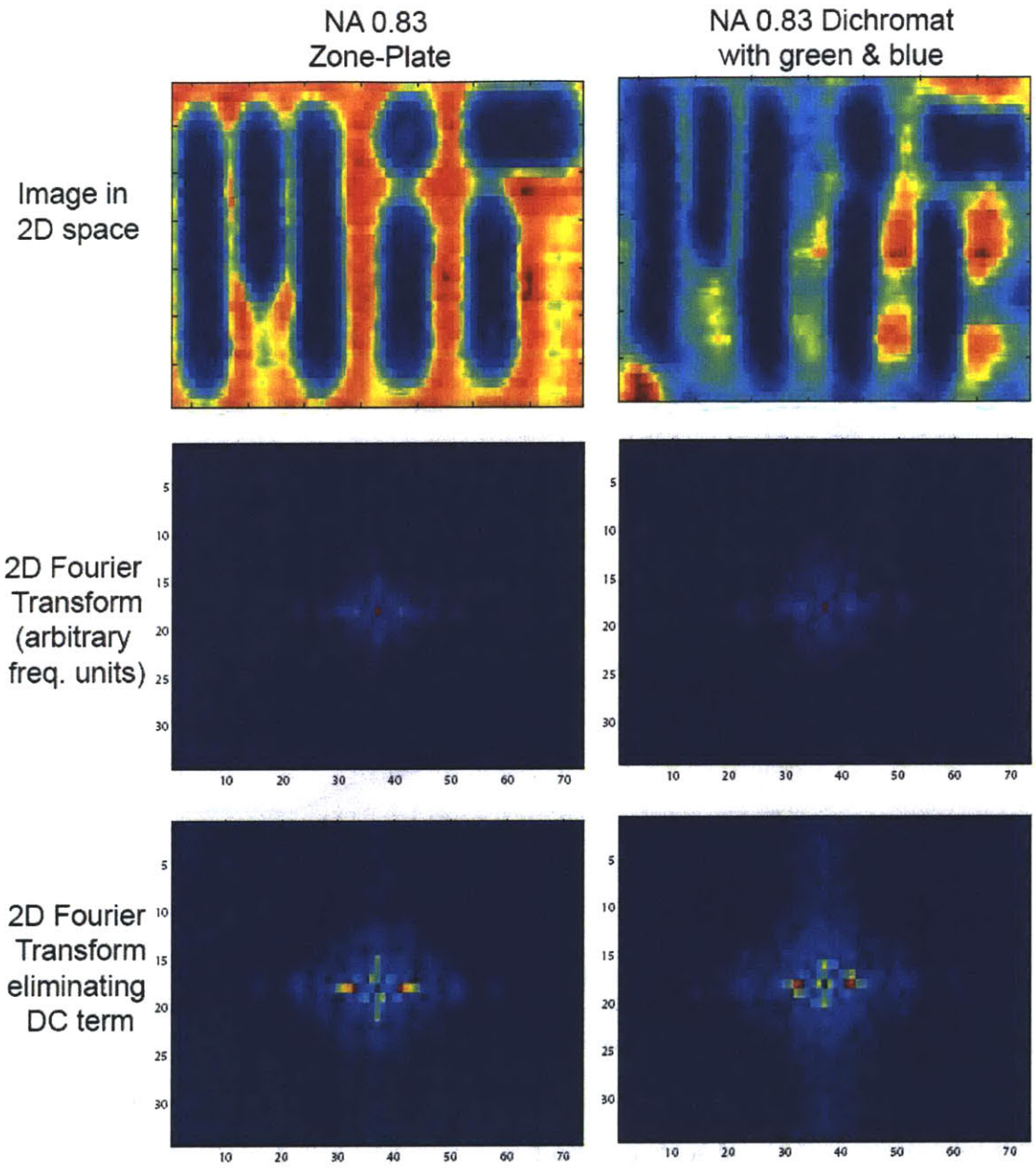

Figure 5-12: Two-dimensional Fourier transform amplitude plots with and without the DC term for images in Fig. 5-11 (b) and (d). Each Fourier transform plot is normalized (in color code) to its maximum value and zero frequency is at the center of each plot. Qualitatively, the Fourier transform plot of the NA = 0.83 dichromat image, obtained with both $\lambda_1 = 405$ nm and $\lambda_2 = 532$ nm illumination, contains a wider spread of frequency components compared to that of the NA = 0.83 Fresnel zone plate image. This is consistent with qualitative visual observations that absorbance modulation images contain higher spatial frequencies.

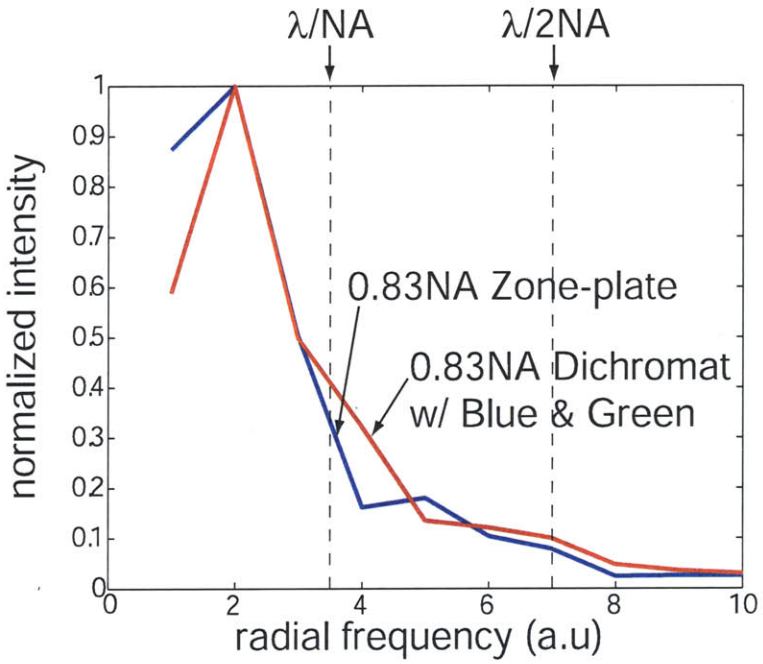

Figure 5-13: Fourier-transform amplitude distribution with respect to radial frequency for Fig. 5-11(b) and Fig. 5-11(d). The Fourier transform amplitudes at each radial frequency is obtained by averaging points on the 2D Fourier transform plots with the same radial frequency. The amplitude curves are normalized to their maximum values and the zero frequency components are removed before normalization. This plot shows that the absorbance modulation image (Fig. 5-11(d)) contains more of the high spatial frequency components than does the zone-plate image (Fig. 5-11(b)).

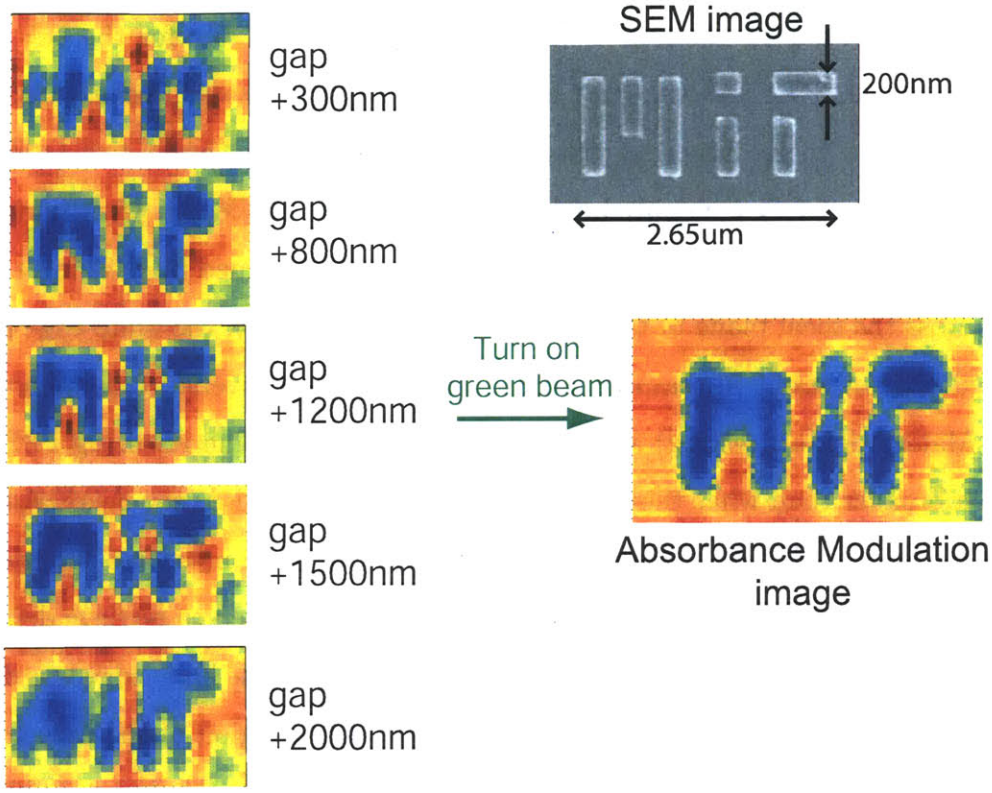

Figure 5-14: Gapping of the imaging system for an NA = 0.83 dichromat. The "gap" here refers to where the reflection signals of the satellite HeNe zone plates are in focus. The series of images taken at different gaps show that gapping is critical to the quality of the image, determined by not only the image contrast, but also the legibility of the logo, and that the optimal gap differs from the focal distances of the satellite zone plates.

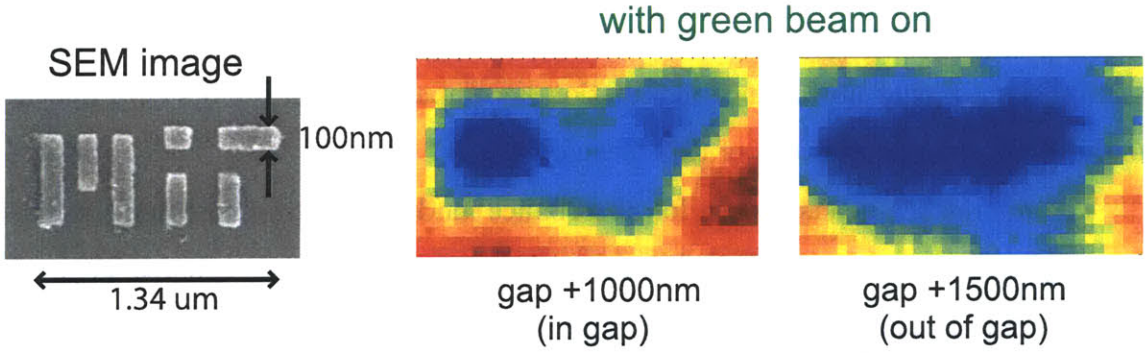

Figure 5-15: Absorbance modulation images obtained with an NA = 0.83 dichromat of an MIT logo with 100-nm linewidth.

be very small in theory. One thing we could tell about the node quality from the PSF characterization in Chapter 4 was an upper bound of 10% for the NA = 0.83 dichromat and 20% for the NA = 0.7 and NA = 0.55 dichromats. Because this upper bound is far too coarse according to simulation results in Chapter two (10% noise in the node completely eliminated absorbance modulation effects), we attempted to measure the quality of the node by scanning a knife edge formed by cleaving a silicon wafer across the focal plane of the dichromats. Unfortunately, the knife-edge profiles obtained did not result in reasonable PSFs after applying inverse Radon transforms (using Matlab default setting) assuming radial symmetry. The resulting PSFs were all much broader than what they were supposed to be and were very noisy. Although inverse Radon transforms intrinsically amplify high frequency noise and suppress DC-terms in the PSF, the knife edge scan method has been applied to retrieve doughnut-shaped profiles in the literature[129]. In this specific paper, the knife edge was directly deposited on the photo-diode, covering half of the active region of the detector. We could not afford to do so with the CCD cameras we have, but this is a possibility for PSF characterizations in the future.

5.4.1 Fluorescent Beads

Another type of resolution standards that is commonly used in the field of super-resolution imaging is fluorescent beads. Although part of the reason is that most other super-resolution imaging techniques rely on fluorescence, fluorescence imaging by itself is interesting and key to many imaging applications in biology. In this section, I show that absorbance modulation is compatible with fluorescence imaging and discuss some limitations of this technique.

The fluorescent beads chosen for absorbance modulation are fluorescent polymer microspheres of sizes 60 nm and 530 nm with "plum purple" fluorophores inside the spheres (Bangs Laboratories, FS02F/2598 and FS03F/7463.) The absorption and emission spectra of the plum purple fluorophores are shown in Fig. 5-16(a). The sizes of the beads were limited by the microsphere supply from Bangs Laboratories. A band pass filter (Chroma Technology Corp. ET460/50m) was placed before the CCD

to filter out both $\lambda_1 = 405$ nm and $\lambda_2 = 532$ nm illuminations and collect only the fluorescence signals. The transmission spectrum of the band pass filter is shown in Fig. 5-17.

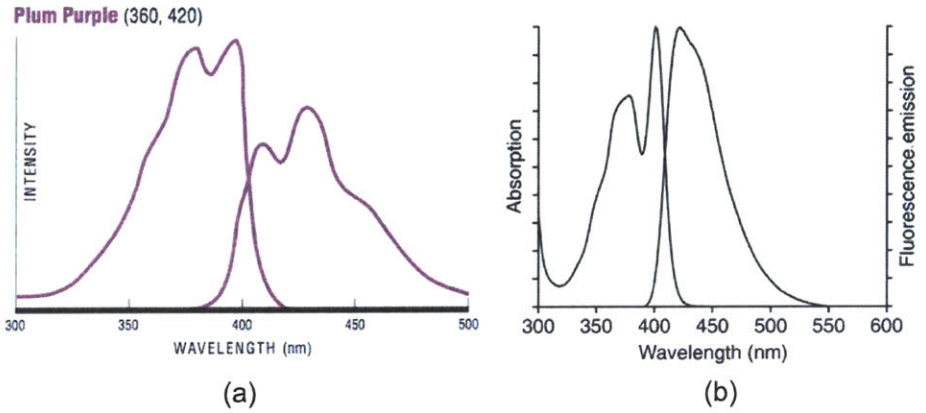

Figure 5-16: Absorption and emission spectra of (a) plum purple fluorophores from Bangs Laboratories and (b) Alexa Flour 405 from Invitrogen (molecular probes).

The fluorescent beads were spin-coated on a two-inch optical flat and dried in air. The sample was first imaged without the AML over-coating for a reference of the brightness and robustness of the fluorescent beads. In general, fluorescence samples are much less robust compared to semiconductor or metal samples and the brightness of the fluorophores can be affected by environmental conditions. The fluorescent beads used in this section contained fluorescent molecules, and so were brighter and less prone to environmental effects. When I tested another fluorophore, Alexa Flour 405, by conjugating the fluorophore on the external surface of latex beads, the fluorophore appeared dim (even without AML overcoating) although Alexa Flour 405 (spectra shown in Fig. 5-16(b)) is known to be a very bright fluorescent molecule. In addition, the fluorescent signal lasted for no more than 30 seconds upon dichromat focused λ_1 illuminations with an input power of 51.9 mW, indicating rapid photo-bleaching at that incident power. The plum purple fluorescent beads were bright and withstood multiple scans with input λ_1 powers as high as 10 mW (photo-bleaching is also observed after many additional scans at this power.) A 23 nm of PVA was spin-coated over the fluorescent beads before spin-coating 220 nm of azobenzene AML

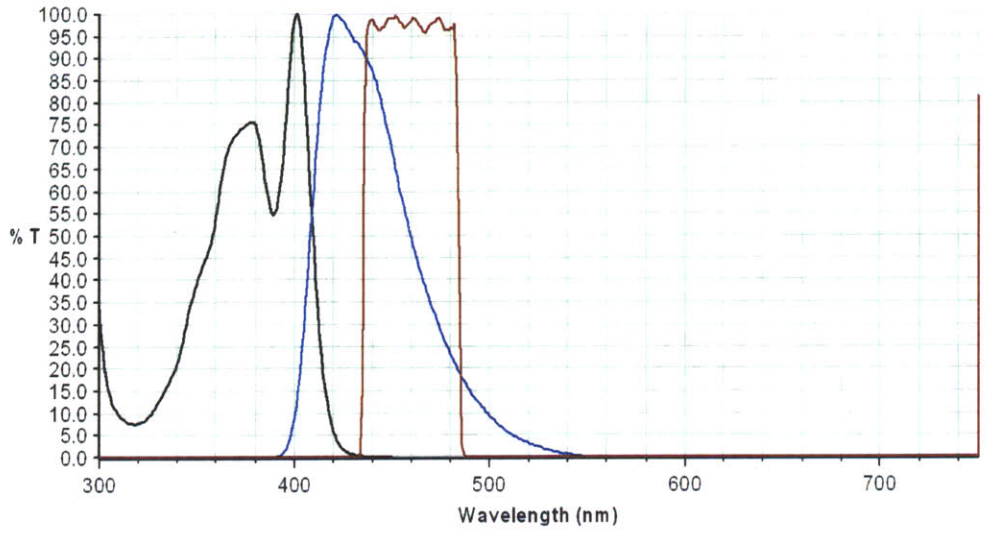

Figure 5-17: Transmission spectrum of the band-pass filter from Chroma Technology Corp. The red curve is the ET460/50m band pass filter used in our experiment to filter out both $\lambda_1 = 405$ nm and $\lambda_2 = 532$ nm illuminations. The black and blue curves are the absorption and emission spectra of Alexa Flour 405.

to protect the beads from being dissolved by the solvent of the AML.

Figure 5-18 shows an absorbance modulation image of the 500 nm plum purple beads with an NA = 0.7 dichromat. When applying $\lambda_2 = 532$ nm illumination in Fig. 5-18(b), the shapes of the beads became more confined and gaps between bead became clearer, but the effect was not as dramatic as for the gold nano-particle samples. One reason may be that the 220-nm AML was not covering the 500-nm fluorescent beads properly. Even if the 220-nm AML formed a conformal coating on the beads, the vertical thickness of the AML becomes effectively thicker at the edges of the beads, complicating image interpretation. When scaling down the λ_1 beam power to achieve higher I_2/I_1 ratios, as shown in Fig. 5-19, we see that the signal to noise level dropped rapidly before observing significant improvements in image resolution. The step-sizes in Fig. 5-18 and Fig. 5-19 were 100 nm and no interpolation was applied to the raw data.

In summary, we demonstrated that absorbance modulation can be made compatible with fluorescent imaging and showed some level of image improvement when applying absorbance modulation. Signal generation and collection efficiency is impor-

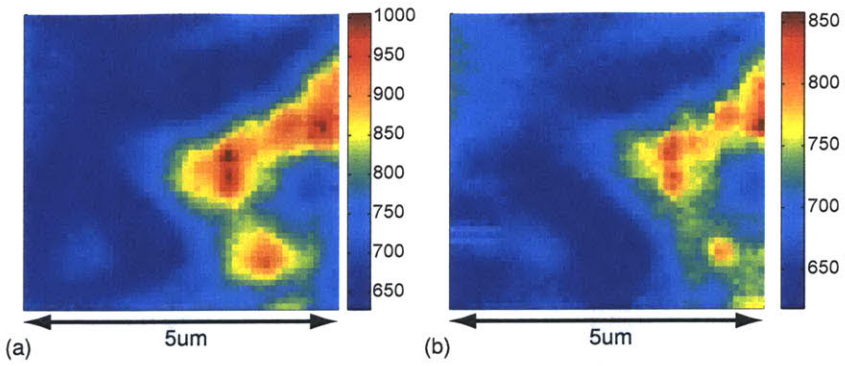

Figure 5-18: Absorbance modulation images of the 500-nm plum purple beads with an NA = 0.7 dichromat. (a) was obtained with the AML coating, but only λ_1 illumination. The power of the input beam of 9.35 mW. (b) was obtained with both λ_1 and λ_2 illuminations at a incident power ratio of 0.74 with the same λ_1 power as in (a).

tant in the design of the system as each fluorophore can only emit a certain number of photons (on average) before they photo-bleach. Because the design of absorbance modulation assumed a flat sample on which to generate highly confined probes, the sample needs to be at least locally flat for absorbance modulation to behave properly. As a result, imaging thick biological samples can be problematic.

5.5 Summary and Future Work

Absorbance modulation provides the means to perform near-field scanning optical lithography and imaging without bringing a physical nanoscale probe in close proximity to the sample. Absorbance modulation can be massively parallelized to enhance throughput for high-speed lithography and imaging. For imaging applications, the technique does not rely on fluorescence and utilizes low light intensities. However, it is clearly limited to surface (2D) imaging and requires that the sample to be relatively flat. In this chapter, we described preliminary results that indicate PSF compression with absorbance modulation to various levels. The practically achievable resolution is limited by the thermal instability of the cis (transparent) isomer of the azobenzene and the quality of 2D node generation at the λ_2 wavelength.

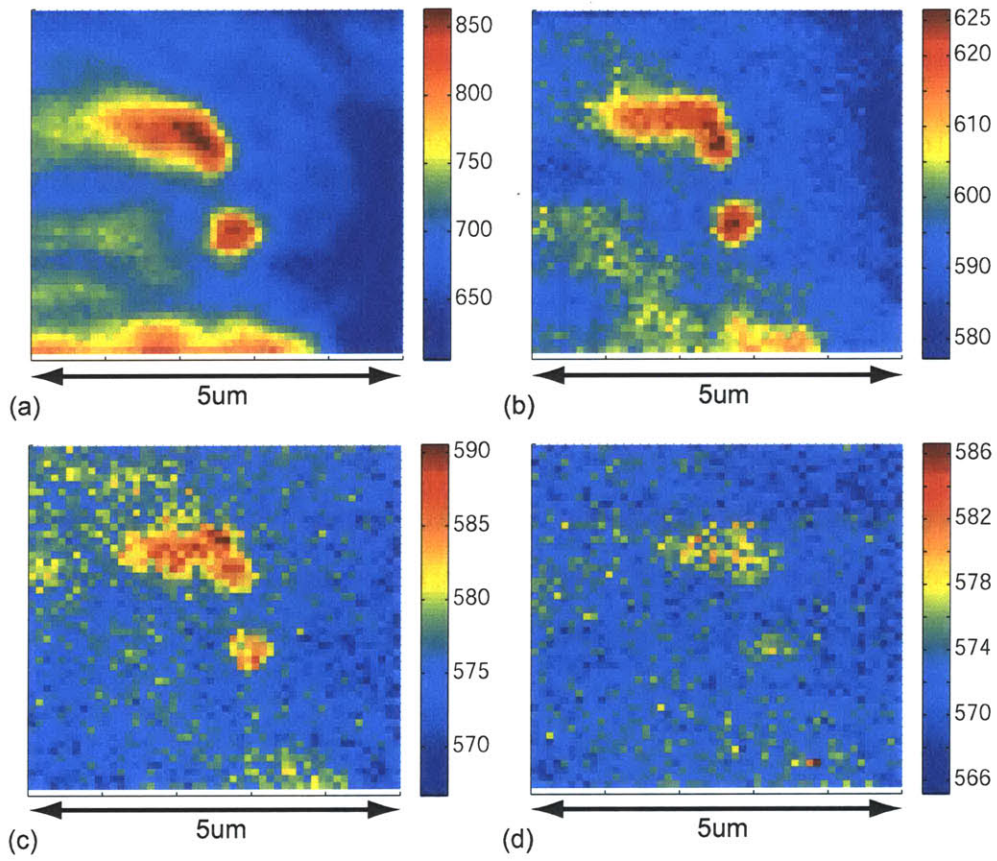

Figure 5-19: Absorbance modulation images of the 500-nm plum purple beads with an NA = 0.7 dichromat at four power ratios: (a) λ_1 power = 9.35 mW, power ratio = 0.74, (b) λ_1 power = 2.79 mW, power ratio = 2.5, (c) λ_1 power = 0.962 mW, power ratio = 7.2, (d) λ_1 power = 0.27 mW, power ratio = 25.6. The λ_2 power was 6.9 mW in all cases.

5.5.1 Alternative Methods for Optical Null Generation

Since the quality of the optical null is crucial to STED-like techniques, many methods have been investigated for the generation of doughnut-shaped illumination profiles. As described earlier in this thesis, doughnut-shaped profiles can be generated by passing a plane wave through a spiral-phase plate or spiral-zone plate. A simplified version utilizes only two π-phase steps and two polarizing beam splitters was used in one of the demonstrations of STED [52]. A circularly or linearly polarized beam is first equally separated into its s- and p- polarized components with a polarizing beam splitter. Two π-phase plates are then inserted separately with the edge of the step parallel to the direction of polarization to create two dark lines. By recombining the two polarizations with another polarizing beam splitter, an optical null at the cross of the two lines can be formed. Doughnut-shaped profiles can also be implemented with more complicated configurations, such as with a spatial-light modulator to impose the required spiral phase [104], with off-axis holograms, or by focusing light beams into a Gauss-Laguerre (doughnut-shaped) mode [44]. Although all above method requires an additional alignment step of the doughnut-shaped deactivation spot with the round-shaped activation spot and can become tricky when massive parallelism is desired, the generally better null quality and higher focusing efficiencies compared to diffractive-optical elements are desirable when scanning only a single pair of activation and deactivation spots.

In some preliminary experiments, we implemented null generation with two phase steps, characterized quality of the null, and tested the null for absorbance modulation imaging. Although photoresist exposures confirmed that there were indeed two line-shaped nodes generated, even aligning those two spots was difficult, not to mention aligning the cross-section of the twos deactivation spots with the round-shaped activation spot. A single-path STED microscope whose beams are self-aligned using a phase plate that selectively modulates the STED beam without affecting the excitation beam showed similar performance in resolution while gaining insensitivity to mechanical drift [130].

5.5.2 Absorbance Modulation Imaging in Reflection Mode

Imaging in reflection mode enables highest collection NA possible because it avoids imaging through the transparent substrate and, therefore, greatly reduces the required minimum working distance of the signal collection objective lens. Imaging in reflection can also bring flexibility in the choice sample substrates in terms of transparency and thickness. Nevertheless, AMI in reflection mode suffers from reduced signal to noise ratio due to having the signal pass through the AML twice. In the case of fluorescence imaging, where the signal wavelength is different from the excitation wavelength, it is possible to select an AML that is less absorptive at the signal wavelength. Preliminary experiments of AMI in reflection with chrome-on-glass resolution standards showed that the contrast curve (or modulation transfer function) is improved when illuminating with both λ_1 and λ_2 light compared to only with the λ_1 illumination, but the image contrast in both cases are low compared to AMI in transmission.

5.5.3 Faster Scanning for Higher Throughput

In many scanning optical microscopes, it is the optical illuminations that are scanned with respect to the sample, not the sample stage. Because steering of optical beams can achieve much higher speeds than translating a physical stage, many single-probe scanning imaging techniques, such as two-photon imaging, confocal imaging, and STED [131], applies this scanning scheme. One important reason why we must scan the stage in the optical setup of this thesis is because of the use of diffractive-optical elements. Planar diffractive-optical elements are ideal for fabrication of a large array of such optics, but only work properly with normally incident illuminations. As a result, deflecting the incident beam will not give the desired effect of scanning the on-axis PSFs. Since scanning with a large array of probes does not reduce scanning time for each field (the minimum field size each probe needs to scan over is determined by the spacing of the probes), when only local imaging over small areas are concerned, and AMI system deploying conventional optics with beam steering could

be advantageous.

Chapter 6

Conclusion

The research described in this thesis was established on previous research in the field of super-resolution optical lithography and optical imaging. With the promising potential of absorbance modulation systems, absorbance-modulation optical lithography (AMOL) and absorbance-modulation imaging (AMI) techniques demonstrated resolution capabilities beyond the optical diffraction limit. Absorbance modulation uses an optical null in conjunction with reversible photochemistry in the absorbance modulation layer (AML). Numerical models of the photo-responses in the AML accessed both transient and steady-state responses of the AML and predicted point-spread function (PSF) scaling to as small as 50-nm full-width at half-maximum (FWHM). Therefore, absorbance modulation provides the means to perform near-field scanning optical lithography or imaging without bringing a physical nanoscale aperture in close proximity to the sample. The practically achievable resolution with absorbance modulation is limited by thermal instability of the AML, quality of the null illumination, and highest I1 to I2 intensity available. A thermally stable AML enabled line exposures as narrow as 40 nm in photoresists with an interference lithography setup.

In order to create the required doughnut-shaped illumination at one wavelength overlapping with round-shaped spots at another wavelength for creation of point apertures in the AML, a binary diffractive-optical element was designed using genetic algorithms and fabricated using scanning-electron-beam lithography. Lithography and imaging results applying this diffractive optical element showed evidence of

point-spread function compression in lithography and contrast enhancement in imaging. Preliminary imaging results of randomly dispersed gold nano-particles indicated imaging of structures spaced by distances as small as one fifths of the illumination wavelength.

Absorbance modulation has several unique properties that distinguish it from existing techniques. Absorbance modulation utilizes low light intensities, which enable parallelization with the potential for high-speed lithography and imaging. With further research on AML materials, it is possible to achieve efficient nano-scale lithography of arbitrary patterns and also surface nano-microscopy at high throughput with resolution far beyond the optical diffraction limit. Improvement on the quality of optical nulls in the doughnut-shape spot and the speed of scanning the spots across the field would further enhance resolution and throughput of absorbance modulation systems. Since absorbance modulation generates near-field apertures independently of the substrate below, this technique does not limit the type of photoresists used for lithography and does not rely on fluorescence for imaging. This attribute enables flexibility in separately optimizing the photoresist and the AML and empowers AMI for studying a variety of photo-responses, including plasmonics and fluorescence.

Appendix A

Absorbance Modulation Simulation Matlab Implimentation Code

A.1 Differential Equations Solver for Steady-State Response in the Absorbance Modulation Layer

A.1.1 The Main Function

```
%%%% Main function to solve coupled ODE
%%%% to calculate beam propagation through thick AML
%%%% Photostationary state is assumed.
%%% Goal is to solve the problem :
%%%% dI1/dz = -alpha1*I1
%%%% dI2/dz = -alpha2*I2
%%% where alpha1(I1, I2, z) and alpha2(I1, I2, z)
%%% Version 2 (RM) updated to simulate with ZP and SPP
%%% Version 3 (RM) updated to loop through a range of Ratios
%%% Also the correct units for intensities used (mols of photons/m^2).
%%% Otherwise the units with kBA wont be correct.
%%% Version 4: include the actual incident power densities on the optic
```

```
%%% This also takes into account the focusing efficiency of the optic.
%%% April 22, 2008: All parameters adjusted for dichromat & azobenzene.
%%% March, 2010: simulated FWHM and Efficiency scaling with 200nm-azo
%%% AML with dichromats (40 zones, 80 zones, 115(160) zones)

% close all;
% clear all;

%%%% constants
Na = 6.022e23; h = 6.626e-34; c = 3e8;
l1 = 400e-9;
l2 = 532e-9;
moles1 = Na*h*c/l1;
moles2 = Na*h*c/l2; % moles of photons at lambda2

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%% Parameters to adjust %%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

Ratio = [2:2:10 20:10:200 250:50:500 600:100:2000 3000:1000:20000];
% ratio of max(I2)/max(I1)
% Ratio=20000; % used for plotting best ratio

I_in1 = 10;
% Watts/m^2 at lambda1 (uniform intensity, incident on the optic)
I_in2 = 10; % Watts/m^2 at lambda2 % 10 W/m^2 is about correct
zmax = 200e-9; % thickness of AML layer unit=m

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
read_from_file =0;
% set to 0 to use analytical lens focal spot & SPP PSFs.
% set to 1 to read from appropriate dichromat PSF file
% set to 2 use standing wave in lambda2 & flood illumination at lambda1
% set to 3 for two standing waves at lambda1 and lambda2 (phase shifted by pi)

if read_from_file == 0
%%% space dimensions in meters
    Rmax = 750e-9;
    dr = 10e-9; r = 0:dr:Rmax; % radial co-ordinate
    rsub= 0:dr:300e-9;
%%% Incident profiles
    NA=0.95; %% Change NA here
    R_D = 40e-6*tan(asin(NA));
% Radius of the dichromat for 0.7NA (80 zones), it is ~40microns.

    I10_t = besselj(1,2*pi*NA/l1*r)./r;
    I10_t(1)=pi*NA/l1;
    I10_t = I10_t/max(I10_t);
    I10_t=I10_t.*I10_t;
    I_scale1 = I_in1*pi*R_D^2/energy2_PSF(r,I10_t,0,max(r));
    I100 = I_scale1*I10_t;

    % Optical Node ~ 2pi/k^2 integral(xJ1(x)dx, 0 to ka)
    for ri= 2: length(r)
    I20_t(ri) = energy2_PSF(2*pi*NA/l2*r,besselj(1,2*pi*NA/l2*r),0,
2*pi*NA/l2*r(ri));
    end
    I20_t(1)=0;
    I20_t = I20_t/max(I20_t);
```

```
    I20_t=I20_t.*I20_t;
    I_scale2 = I_in2*pi*R_D^2/energy2_PSF(r,I20_t,0,max(r));
    I200 = I_scale2*(I20_t + 0.001); % Can add noise to doughnut beam

elseif read_from_file == 1 % dichromat
    filename='PSFs_80.txt';
    data = load(filename);
    r_t = data(1:20,1)*l1; I10_t = data(1:20,2); I20_t=data(1:20, 3); clear data;
    I10_t = I10_t/max(I10_t); I20_t = I20_t/max(I20_t);

    % R_D = 60e-6; % Radius of 0.83NA (160 zones) dichromat, it is ~60microns.
     R_D = 39.198e-6; % Radius of 0.7NA (80 zones) dichromat, it is ~40microns.
    % R_D = 26.5e-6; % Radius of 0.55NA (40 zones) dichromat, it is ~26.5microns.

    I_scale1 = I_in1*pi*R_D^2/energy_PSF(r_t,I10_t,max(r_t))*1/moles1;
    I10_t = I_scale1*I10_t;
    I_scale2 = I_in2*pi*R_D^2/energy_PSF(r_t,I20_t,max(r_t))*1/moles2;
    I20_t = I_scale2*I20_t;

    dr = 10e-9; r = 0:dr:max(r_t);
    I100 = interp1(r_t,I10_t,r,'cubic'); I200 = interp1(r_t,I20_t,r,'cubic');
    I200 = I200 + 0.0*max(I200); % Can add noise to doughnut beam

elseif read_from_file == 2
%%% use standing wave in lambda2 & flood illumination at lambda1
    P2 = 0.345e-6; % period of standing wave at lambda2
    dr = 5e-9; r = 0:dr:P2;
    I_scale2 = I_in2*1/moles2;
    I20 = 2*I_scale2*sin(2*pi/(P2*2)*r).^2; I20 = I20 + 0.0*max(I20);
    I_scale1 = I_in1*1/moles1;
```

```
    I100 = I_scale1*ones(size(I20));

elseif read_from_file == 3  %%% use standing waves at both lambda2 & at lambda1
    P2 = 0.35e-6; % period of standing wave at lambda2
    dr = 10e-9; r = 0:dr:5*P2;
    I_scale2 = I_in2*1/moles2;
    I20 = 2*I_scale2*sin(2*pi/(P2*2)*r).^2; I20 = I20 + 0.0*max(I20);
    I_scale1 = I_in1*1/moles1;
    P1 = 0.176e-6; %  period of standing wave at lambda1
    I100 = 2*I_scale1*cos(2*pi/(P1*2)*r).^2; I100 = I100 + 0.0*max(I20);
end

% figure(1); plot(r*1e9,I10/max(I10),'b',r*1e9,I20/max(I20),'r');
% xlabel('r(nm)'); ylabel('Normalized intensity');

%%%% choose photochrome %%%%
photochrome = 1;

%%% Photochromic parameters
if photochrome == 1 % azobenzene polymer
    e1a=30000;
    e1b=12500;
    e2a=1250;
    e2b=3000;
    A0 =400.024; % mol/m^3
    Q1ab=.0037;
    Q1ba=.0021;
    Q2ab=.0021;
    Q2ba=.0072;
    Kba=.002;
```

```
elseif photochrome == 2 % thienylethene at 325nm & 633nm
    e1a=32675;
    e1b=11004;
    e2a=157.81;
    e2b=18133;
    A0 =306.044*2/2; % mol/m^3 for initial absorbance of 2 with 200nm thickness
    Q1ab=0.2;
    Q1ba=9.6e-4;
    Q2ab=Q1ab;
    Q2ba=Q1ba;
    Kba=0;
end

dz = 5e-9; % 5nm step in z
z_vector = 0:dz:zmax;

myfun = @(t, y) dIdz(t, y,e1a, e1b, e2a, e2b, Q1ab, Q2ab, Q1ba, Q2ba, Kba, A0);
options = odeset('RelTol',1e-4);

I1_matrix = zeros(length(z_vector), length(r));
I2_matrix = zeros(size(I1_matrix));
A_matrix = zeros(size(I1_matrix));

FWHM = zeros(size(Ratio));
half_FWHM_in = interp1(I100/max(I100),r,0.5,'cubic');
Eff = zeros(size(Ratio));
Energy_in = energy_PSF(r,I100, half_FWHM_in);
Peak_in = I100(1);
```

```
for cnt_ratio = 1:length(Ratio) %%% Loop through different ratios
%%% Fix I1 intensity and scale I2 up %%%
    I10 = I100;
    I20 = I200*Ratio(cnt_ratio);

%%% Fix I2 intensity and scale I1 down %%%
%     I10 = I100/Ratio(cnt_ratio);
%     I20 = I200;
%Ratio(cnt_ratio)
    for cntr = 1:length(r)%%
    % Z is vector of z values returned by ode45
    % II is a matrix, where II(:,1) is I1 as a function of z
    % and II(:,2) is I2 as a function of z
        [Z, II] =  ode15s(myfun, z_vector, [I10(cntr) I20(cntr)], options);
    %%% ode15s is very fast
        I1_matrix(:, cntr) = II(:,1);
        I2_matrix(:, cntr) = II(:,2);
    end

    FWHM(cnt_ratio) =
2*interp1(I1_matrix(length(z_vector),:)/max(I1_matrix(length(z_vector),:)),r,
0.5,'cubic');
    Energy_out = energy_PSF(r, I1_matrix(length(z_vector),:), FWHM(cnt_ratio)/2);
    Eff_peak(cnt_ratio) = I1_matrix(length(z_vector),1)/Peak_in;
    Eff(cnt_ratio) = Energy_out/Energy_in;

%%% Plot cross sectional profiles when plot_cross ==1 %%%
    plot_cross =0;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    if plot_cross ==1
```

```
        figure;
        subplot(3,1,1);
        imagesc(r*1e9, z_vector*1e9, I1_matrix); title('I_1'); colorbar;
        xlabel('r(nm)'); ylabel('z(nm)');
        axis([0,750,0,200])

        subplot(3,1,2);
        imagesc(r*1e9, z_vector*1e9, I2_matrix); title('I_2'); colorbar;
        xlabel('r(nm)'); ylabel('z(nm)');
        axis([0,750,0,200])

        subplot(3,1,3);
        A_matrix = A0*(Q1ba*e1b*I1_matrix + Q2ba*e2b*I2_matrix + Kba)
./((Q1ab*e1a+Q1ba*e1b)*I1_matrix + (Q2ab*e2a+Q2ba*e2b)*I2_matrix + Kba );
        imagesc(r*1e9, z_vector*1e9, A_matrix); title('[A]'); colorbar;
        xlabel('r(nm)'); ylabel('z(nm)');
        axis([0,750,0,200])

        figure;
        plot(r*1e9, I10/max(I10), r*1e9, I20/max(I20), r*1e9,
I1_matrix(length(z_vector),:)/max(I1_matrix(length(z_vector),:)));
        title(['Azo-AML,incident I_1= ',num2str(I_in1),' W/m^2,I_2/I_1= ',
num2str(Ratio(cnt_ratio)),', input = SPP NA ',num2str(NA)])
        xlabel('r(nm)')
        ylabel('Normalized Intensity')
        legend('I_{1}','I_{2}','I_{out}')
    end
end

%%% plot trend when plot_trend==1 %%%
```

```
plot_trend=1;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if plot_trend == 1

    Peak_ratio= Ratio*I_scale2/I_scale1;
    I_scale2/I_scale1;

%%% Energy Efficiencies %%%
%    figure; plot(Ratio,FWHM*1e9);
%    title(['FWHM vs incident I_2/I_1 for Azo AML input = ',filename])
%    xlabel(['incident I_2/I_1 ratio (fix I_2= ',num2str(I_in2),' W/m^2)'])
%    ylabel('Full-Width-Half Maximum (nm)')

%    figure; hold on; plot(Peak_ratio,FWHM*1e9);
%    title('FWHM over focal-plane I_{2}/I_{1} ratio for Azo AML')
%    xlabel('I_{2}/I_{1} peak intensity ratio (fix I_{2})')
%    ylabel('Full-Width-Half Maximum (nm)')

    figure; semilogx(Ratio,FWHM*1e9);
    title(['FWHM vs incident I_2/I_1 for Azo AML input = SPP NA ',num2str(NA)])
    xlabel(['incident I_2/I_1 ratio (fix I_1= ',num2str(I_in1),' W/m^2)'])
    ylabel('Full-Width-Half Maximum (nm)')

%    figure; semilogx(Peak_ratio,FWHM*1e9);
%    title('FWHM over focal-plane I_{2}/I_{1} ratio for Azo AML')
%    xlabel('I_{2}/I_{1} peak intensity ratio (fix I_{2})')
%    ylabel('Full-Width-Half Maximum (nm)')

%    figure; semilogx(Ratio,Eff);
%    title(['Energy Eff. vs incident I_2/I_1 for Azo AML input = ',filename])
```

```
%    xlabel(['incident I_2/I_1 ratio (fix I_2= ',num2str(I_in2),' W/m^2)'])
%    ylabel('Energy Efficiency')

%    figure; hold on; semilogx(Peak_ratio,Eff);
%    title('Energy Eff. vs focal-plane I_{2}/I_{1} ratio for Azo AML')
%    xlabel('I_{2}/I_{1} peak intensity ratio (fix I_{2})')
%    ylabel('Energy Efficiency')

    figure; loglog(Ratio,Eff);
    title(['Energy Eff. vs incident I_2/I_1, Azo AML input=SPP NA ',num2str(NA)])
    xlabel(['incident I_2/I_1 ratio (fix I_1= ',num2str(I_in1),' W/m^2)'])
    ylabel('Energy Efficiency')

%    figure; loglog(Peak_ratio,Eff);
%    title('Energy Eff. vs focal-plane I_{2}/I_{1} ratio for Azo AML')
%    xlabel('I_{2}/I_{1} peak intensity ratio (fix I_{2})')
%    ylabel('Energy Efficiency')

%%% Peak Efficiencies %%%
%    figure; semilogx(Ratio,Eff_peak);
%    title(['Peak Intensity Eff. vs incident I_2/I_1, Azo AML input= ',filename])
%    xlabel(['incident I_2/I_1 ratio (fix I_2= ',num2str(I_in2),' W/m^2)'])
%    ylabel('Peak Intensity Transmission')

%    figure; loglog(Peak_ratio,Eff_peak);
%    title('Peak Intensity Eff. vs focal-plane I_{2}/I_{1} ratio for Azo AML')
%    xlabel('I_{2}/I_{1} peak intensity ratio (fix I_{2})')
%    ylabel('Peak Intensity Transmission')

    figure; loglog(Ratio,Eff_peak);
```

```
    title(['Peak Int. Eff. vs incident I_2/I_1,Azo AML input=SPP NA',num2str(NA)])
    xlabel(['incident I_2/I_1 ratio (fix I_1= ',num2str(I_in1),' W/m^2)'])
    ylabel('Peak Intensity Transmission')

    figure; plot(r*1e9, I10/max(I10), r*1e9, I20/max(I20), r*1e9,
I1_matrix(length(z_vector),:)/max(I1_matrix(length(z_vector),:)));
    title(['Azo-AML,incident I_1= ',num2str(I_in1),' W/m^2,I_2/I_1= ',
num2str(Ratio(length(Ratio))),', input = SPP NA ',num2str(NA)])
    xlabel('r(nm)')
    ylabel('Normalized Intensity')
    legend('I_{1}','I_{2}','I_{out}')
end

%%% Note: Below are for only the last member of the vector: Ratio %%%
save_values_file = 0; % set to 1 to save Ratio values to a file.
%%%%%%%%%%%%%%%%%%%%

if save_values_file == 1
    filename2 = '20100222PSFs_80_FWHM_Eff.txt';
%%% enter appropriate filenmae here to save ratio values
    fid = fopen(filename2,'w');
    fprintf(fid,'%s %s %s %s',['Incident_Ratio ','Focal_Plane_Ratio ',
'FWHM(nm) ','Efficiency']);
    fprintf(fid, '\n'); %%% TO FILE %%%
    fprintf(fid, '%g %g %g %g\n', [Ratio; Peak_ratio; FWHM*1e9; Eff]);
    fclose(fid);
    disp('File saved as:');
    disp(filename2);
end
```

A.1.2 Point-Spread-Function Energy Calculation

Integrate from a to b

```
function y  = energy2_PSF(r,I,a,b)
%%% integrate I*r*dr from a to b.
    y = quadv(@PSF_integrand,a,b);
    function w = PSF_integrand(x)
    %%% interpolating the integrand.
       w = interp1(r, I, x, 'cubic').*x;
    end
end
```

Integrate from 0 to a

```
function y  = energy_PSF(r,I,a)
% integrate I*r*dr from 0 to a.
    y = quadv(@PSF_integrand,0, a);
    function w = PSF_integrand(x)
    %%% interpolating the integrand.
        w = interp1(r, I, x, 'cubic').*x;
    end
end
```

A.1.3 System of Differential Equations

```
function dy = dIdz(z,y, e1a,e1b,e2a,e2b, Q1ab,Q2ab,Q1ba,Q2ba, Kba, A0)
%%% setup system of differential equations
%%% y  = [I1; I2]
    dy = zeros(2, 1);
    A = A0*(Q1ba*e1b*y(1) + Q2ba*e2b*y(2) + Kba)
/((Q1ab*e1a+Q1ba*e1b)*y(1) + (Q2ab*e2a+Q2ba*e2b)*y(2) + Kba );
    dy(1) = -log(10)*((e1a-e1b)*A+e1b*A0)*y(1);
```

```
dy(2) = -log(10)*((e2a-e2b)*A+e2b*A0)*y(2);
```

A.2 Finite Difference Solver for Transient Response in the Absorbance Modulation Layer

A.2.1 The Main Function for Transient Response and Intensity Scaling

```
%%%% Main function to solve coupled PDE with finite difference
%%%% to calculate beam propagation through thick AML
%%% April, 2010 (Sidney)

%close all;
clear all;
%hold on;

%%% constants %%%
Na = 6.022e23; h = 6.626e-34; c = 3e8;
l1 = 400e-9; %unit= meter
l2 = 532e-9;
moles1 = Na*h*c/l1; % energy of one mole of photons at lambda2
moles2 = Na*h*c/l2; % energy of one mole of photons at lambda2
Em1 = Na*h*c/l1; % energy of one mole of photons at lambda2
Em2 = Na*h*c/l2; % energy of one mole of photons at lambda2

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%   Parameters to adjust   %%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%aviobj=avifile('N80_ratio_2000_test_2.avi'); %%% to make movie
%%% enter appropriate filename here to save PSF values
```

```
filename1 = 'PSFs_ratio.txt';
%%% enter appropriate filename here to save Ratio trend values
filename2 = '20101220PSFs_SPP80.txt';
% ratio of max(I2)/max(I1)
Ratio = [2:2:10 20:10:40 50:50:200 300:100:500 600:200:2000];
%Ratio=1900; % used for plotting best ratio
NA=0.7; %% Define NA of the dichromat or lens
dr = 10e-9;  % r grid step size (unit = m)

I_in1 = 10; % W/m^2 at lambda1 (uniform intensity on dichromat)
I_in2 = 10; % W/m^2 at lambda2 % 10 W/m^2 is about correct
BG_2=0.00; % Uniform background for lambda2 profile(in ratio to peak)
make_movie=0;

%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Select Input Profile %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%
read_from_file =0;
%%%%%%%%%%%%%%%%
% set to 0 for focused spiral phase plate PSFs
% set to 1 to read from appropriate dichromat PSF file

if read_from_file == 0
% focused plane wave and SPP (charge m=1) output with lens of NA
%%% space dimensions unit is meters
    Rmax = 750e-9;
    r = 0:dr:Rmax; % radial co-ordinate
    rsub= 0:dr:300e-9;
    R_D = 40e-6*tan(asin(NA));
% Radius of 0.7NA (80 zones) dichromat, it is ~40microns.
```

```
%%% Theoretical Incident Profiles %%%
    I10_t = besselj(1,2*pi*NA/l1*r)./r;
    I10_t(1)=pi*NA/l1;
    I10_t = I10_t/max(I10_t);
    I10_t=I10_t.*I10_t;
    I_scale1 = I_in1*pi*R_D^2/energy_PSF(r,I10_t,0,max(r));
    I100 = I_scale1*I10_t;

    % Optical Node ~ 2pi/k^2 integral(xJ1(x)dx, 0 to ka)
    for ri= 2: length(r)
        I20_t(ri) = energy_PSF(2*pi*NA/l2*r,besselj(1,2*pi*NA/l2*r),
0,2*pi*NA/l2*r(ri));
    end
    I20_t(1)=0;
    I20_t = I20_t/max(I20_t);
    I20_t=I20_t.*I20_t;
    I_scale2 = I_in2*pi*R_D^2/energy_PSF(r,I20_t,0,max(r));
    I200 = I_scale2*(I20_t + BG_2*max(I20_t)); % Add uniform background

%    figure; plot(r,I10_t,r,I200/max(I200))
%    title('Normalized input functions for focused round and SPP spot')
%    xlabel('r(nm)'); ylabel('normalized intensity');
%    legend('round spot','SPP spot')

elseif read_from_file == 1 % dichromat
%%% Choose dichromat radius %%%
     R_D = 60e-6; %Radius of 0.83NA (160 zones) dichromat, ~60microns.
    %R_D = 39.2e-6;%Radius of 0.7NA (80 zones) dichromat, ~40microns.
    %R_D = 26.5e-6;%Radius of 0.55NA (40 zones) dichromat,~26.5microns.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
    data = load('PSFs_115.txt');
r_t = data(1:20,1)*l1; I10_t = data(1:20,2); I20_t=data(1:20, 3); clear data;
    I10_t = I10_t/max(I10_t); I20_t = I20_t/max(I20_t);
% Calculate focal spot intensity scaling factor
% based on incident intensity
    I_scale1 = I_in1*pi*R_D^2/energy_PSF(r_t,I10_t,0,max(r_t));
% Sidney:change moles1 scaling to coefficient, so I unit is still W/m^2
    I10_t = I_scale1*I10_t;
    I_scale2 = I_in2*pi*R_D^2/energy_PSF(r_t,I20_t,0,max(r_t));
    I20_t = I_scale2*I20_t;
    r = 0:dr:max(r_t);
    I100 = interp1(r_t,I10_t,r,'cubic');
    I200 = interp1(r_t,I20_t,r,'cubic');
    I200 = I200 + BG_2*max(I200); % Add uniform background
end
%figure(1); plot(r*1e9,I10/max(I10),'b',r*1e9,I20/max(I20),'r');
% xlabel('r(nm)'); ylabel('Normalized intensity');

%%%%%%%%%%%%%%%
photochrome = 1;
%%%%%%%%%%%%%%%
%%% Photochromic parameters
if photochrome == 1 % azobenzene polymer
    e1a=30000*log(10); % all extinction coefficient multiplied by ln(10) to account
    e1b=12500*log(10);
    e2a=1250*log(10);
    e2b=3000*log(10);
    A0 =400.024; % mol/m^3
    Q1ab=.0037;
```

```
    Q1ba=.0021;
    Q2ab=.0021;
    Q2ba=.0072;
    Kba=.002;

elseif photochrome == 2 % thienylethene at 325nm & 633nm
    e1a=32675*log(10);
    e1b=11004*log(10);
    e2a=157.81*log(10);
    e2b=18133*log(10);
    A0 =306.044*2/2;
% mol/m^3 for initial absorbance of 2 with 200nm thickness
    Q1ab=0.2;
    Q1ba=9.6e-4;
    Q2ab=Q1ab;
    Q2ba=Q1ba;
    Kba=0;
end

alpha1A=e1a*Q1ab/Em1;
alpha2A=e2a*Q2ab/Em2;
alpha1B=e1b*Q1ba/Em1;
alpha2B=e2b*Q1ba/Em2;
%%%%%%%%%%%%%%%%%%%%
zmax = 200e-9; % thickness of AML layer unit = meter
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Choose step sizes and simulation ranges %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
dz = 1e-9; % 1nm step in z (unit = m)
z_vector = 0:dz:zmax;
```

```
dt = 0.1e-3; % 0.1 ms step in time t (unit = sec)
tmax = 200e-3; % total simulation time, unit = s
t_vector = 0:dt:tmax;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%% Create empty vectors
I1_matrix = zeros(length(z_vector), length(r));
I2_matrix = zeros(size(I1_matrix));
%A_matrix = zeros(length(z_vector), length(r),length(t_vector));
% record all A matrix evolution, but memory costly
%A_matrix(:,:,1)=A0*ones(length(z_vector), length(r));
% if record all A matrix evolution
A_matrix = A0*ones(length(z_vector), length(r));
A_ctr=zeros(1,length(t_vector));
A_ctr(1)=A_matrix(length(z_vector),1,1);

FWHM = zeros(1,length(t_vector));
Eff_peak = zeros(1,length(t_vector));
%% Peak (r=0) intensity efficiency of PSF
half_FWHM_in = interp1(I100/max(I100),r,0.5,'cubic');
%Eff = zeros(size(Ratio)); %% Energy Efficiency of PSF
%Energy_in = energy_PSF(r,I100, half_FWHM_in);
% Incident energy in FWHM
FWHM_R= zeros(size(Ratio));
Eff_peak_R = zeros(size(Ratio));
% Peak (r=0) intensity efficiency of PSF
Peak_in = I100(1); %% Incident peak intensity (lambda 1)

k=0; % initialize counter for plotting
for cnt_ratio= 1:length(Ratio) %%% Loop through different ratios
```

```
%    tic
    kp=0; % number of transient plots
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% I2 scale up %%%
%%% fix I1 when changing ratio %%%
    I10=I100;
    I20=I200*Ratio(cnt_ratio);
%%% I1 scale down %%%
%%% fix I2 when changing ratio %%%
%    I10 = I100/Ratio(cnt_ratio);
%    I20 = I200;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Initialization for Finite Difference %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    I1_matrix(1,:)=I10;
    I2_matrix(1,:)=I20;
    I1_matrix_s(1,:)=I10;
    I2_matrix_s(1,:)=I20;
    A_matrix = A0*ones(length(z_vector), length(r));
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    for cnt_z = 1:length(z_vector)-1
%%% Initial condition for entire AML volume
I1_matrix(cnt_z+1,:)=(1-dz/2*e1a*A0)/(1+dz/2*e1a*A0)*I1_matrix(cnt_z,:);
I2_matrix(cnt_z+1,:)=(1-dz/2*e2a*A0)/(1+dz/2*e2a*A0)*I2_matrix(cnt_z,:);
    end
        FWHM(1) = 2*interp1(I1_matrix(length(z_vector),:)
/max(I1_matrix(length(z_vector),:)),r,0.5,'cubic');

    for cnt_t = 1:length(t_vector)-1 % step forward A to next time step
```

```
        A_prev=A_matrix;
        A_matrix= (1-dt/2*(I1_matrix*(alpha1A+alpha1B)+
I2_matrix*(alpha2A+alpha2B)+Kba))./(1+dt/2*(I1_matrix*(alpha1A+alpha1B)
+I2_matrix*(alpha2A+alpha2B)+Kba)).*A_prev
+dt*A0.*(I1_matrix*alpha1B+I2_matrix*alpha2B+Kba)./
(1+dt/2*(I1_matrix*(alpha1A+alpha1B)+I2_matrix*(alpha2A+alpha2B)+Kba));
        A_ctr(cnt_t+1)=A_matrix(length(z_vector),1);

        for cnt_z = 1:length(z_vector)-1
%%% Compute I1, I2 profiles through z with given A
I1_matrix(cnt_z+1,:)=(1-dz/2.*((e1a-e1b).*A_matrix(cnt_z,:)+e1b*A0))./
(1+dz/2.*((e1a-e1b).*A_matrix(cnt_z,:)+e1b*A0)).*I1_matrix(cnt_z,:);
I2_matrix(cnt_z+1,:)=(1-dz/2.*((e2a-e2b).*A_matrix(cnt_z,:)+e2b*A0))./
(1+dz/2.*((e2a-e2b).*A_matrix(cnt_z,:)+e2b*A0)).*I2_matrix(cnt_z,:);
        end

        FWHM(cnt_t+1) = 2*interp1(I1_matrix(length(z_vector),:)
/max(I1_matrix(length(z_vector),:)),r,0.5,'cubic');
%FWHM at AML output (z=zmax)
        Eff_peak(cnt_t+1) = I1_matrix(length(z_vector),1)/Peak_in;
%Peak intensity efficiency at AML output (z=zmax)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%% Make Transient Movie %%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        make_movie=0; % set to 1 to make movie
%%%%%%%%%%%%%
        if make_movie==1
            figure(20);
            subplot(3,1,1);
```

```
    imagesc(r*1e9, z_vector*1e9, I1_matrix); title('I_1');
    caxis auto;
    colorbar;
    xlabel('r(nm)'); ylabel('z(nm)');
    axis([0,750,0,200])

    subplot(3,1,2);
    imagesc(r*1e9, z_vector*1e9, I2_matrix); title('I_2');
    caxis auto;
    colorbar;
    xlabel('r(nm)'); ylabel('z(nm)');
    axis([0,750,0,200])

    subplot(3,1,3);
    imagesc(r*1e9, z_vector*1e9, A_matrix); title('[A]');
    caxis([0 400]);
    colorbar;
    xlabel('r(nm)'); ylabel('z(nm)');
    axis([0,750,0,200])

    frame=getframe(figure(20));
    aviobj=addframe(aviobj,frame);
end

k=k+1; % advance counter for plotting
if k==10 && kp <0 % plot every 10 time frames for kp plots
    kp=kp+1;
    figure(21);
    %figure; % used to plot all frames
```

```
            subplot(3,1,1);
            imagesc(r*1e9, z_vector*1e9, I1_matrix);
title(['I1(Ratio=',num2str(Ratio(cnt_ratio)),'),time=',num2str(kp),'ms']);
            caxis auto;
            colorbar;
            xlabel('r(nm)'); ylabel('z(nm)');
            axis([0,750,0,200])

            subplot(3,1,2);
            imagesc(r*1e9, z_vector*1e9, I2_matrix); title('I_2');
            caxis auto;
            colorbar;
            xlabel('r(nm)'); ylabel('z(nm)');
            axis([0,750,0,200])

            subplot(3,1,3);
            imagesc(r*1e9, z_vector*1e9, A_matrix); title('[A]');
            caxis([0 400]);
            colorbar;
            xlabel('r(nm)'); ylabel('z(nm)');
            axis([0,750,0,200])
            k=0; % reset counter for plotting
        end
    end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Plot final/steady-state intensity, concentration profile %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
plot_cross =0; % set to 1 to plot profile crossections along z
%%%%%%%%%%%%%%
```

```
    if plot_cross ==1
        figure;
        subplot(4,1,1);
        imagesc(r*1e9, z_vector*1e9, I1_matrix);
title(['I_1 (Ratio=',num2str(Ratio),', N=80)']);
caxis auto; colorbar;
        xlabel('r(nm)'); ylabel('z(nm)');
        axis([0,750,0,200])
        subplot(4,1,2);
        imagesc(r*1e9, z_vector*1e9, I2_matrix);
title('I_2'); caxis auto; colorbar;
        xlabel('r(nm)'); ylabel('z(nm)');
        axis([0,750,0,200])

        subplot(4,1,3);
        imagesc(r*1e9, z_vector*1e9, A_matrix);
title(['A_{m}(t=',num2str(tmax*1e3),'ms,dt=',num2str(dt*1e6),'us)']);
        caxis([0 400]);
        colorbar;
        xlabel('r(nm)'); ylabel('z(nm)');
        axis([0,750,0,200])

        A_matrix_s = A0*(alpha1B*I1_matrix + alpha2B*I2_matrix + Kba)./
((alpha1A+alpha1B)*I1_matrix+(alpha2A+alpha2B)*I2_matrix + Kba);
        subplot(4,1,4);
        imagesc(r*1e9, z_vector*1e9, A_matrix_s);
title('A_{steady state}');
        caxis([0 400]);
        colorbar;
        xlabel('r(nm)'); ylabel('z(nm)');
```

```
        axis([0,750,0,200])

        for cnt_z = 1:length(z_vector)-1
%%% Compute I1, I2 profiles through z with given A (steady state)
I1_matrix_s(cnt_z+1,:)=(1-dz/2.*((e1a-e1b).*A_matrix_s(cnt_z,:)+e1b*A0))
./(1+dz/2.*((e1a-e1b).*A_matrix_s(cnt_z,:)+e1b*A0))
.*I1_matrix_s(cnt_z,:);
I2_matrix_s(cnt_z+1,:)=(1-dz/2.*((e2a-e2b).*A_matrix_s(cnt_z,:)+e2b*A0))
./(1+dz/2.*((e2a-e2b).*A_matrix_s(cnt_z,:)+e2b*A0))
.*I2_matrix_s(cnt_z,:);
        end

%%% plot input, output intensities and steady state output
        figure;
        plot(r*1e9,I10/max(I10),r*1e9,I20/max(I20),r*1e9,
I1_matrix(length(z_vector),:)/max(I1_matrix(length(z_vector),:)),r*1e9,
I1_matrix_s(length(z_vector),:)/max(I1_matrix_s(length(z_vector),:)));
        title(['Azo-AML=200nm,incident I_{2}=10W/m^{2},I_{2}/I_{1}
Ratio= ',num2str(Ratio)])
        legend('I_{1}','I_{2}','I_{out}','I_{steady}')

        for cnt_z = 1:length(z_vector)
% Compute FWHM for output I1 and steady state I1
            FWHM_cross(cnt_z) = 2*interp1(I1_matrix(cnt_z,:)
/max(I1_matrix(cnt_z,:)),r,0.5,'cubic');
            FWHM_cross_s(cnt_z) = 2*interp1(I1_matrix_s(cnt_z,:)
/max(I1_matrix_s(cnt_z,:)),r,0.5,'cubic');
        end

        figure;
```

```
% plot FWHM vs depth into AML (z) for final I1 and steady state I1
        plot(z_vector*1e9,FWHM_cross*1e9,z_vector*1e9,FWHM_cross_s*1e9);
        title(['FWHM over AML thickness for peak ratio= ',
num2str(Ratio),', I_{2}=10W/m^{2}'])
        xlabel('z(nm)')
        ylabel('Full-Width-Half Maximum (nm)')
        legend('FWHM_{final}','FWHM_{steady}')
    end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%% Plot transient w.r.t. time %%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    plot_transient =0; % set to 1 to plot time transient
%%%%%%%%%%
    if plot_transient ==1
% plot time transient for central concentration (at \rho=0)
        figure;
        plot(t_vector,A_ctr)
        xlabel('time (s)')
        ylabel('A matrix')
        title(['Central (r=0) absorbance in AML over time for peak
ratio= ',num2str(Ratio),', I_{2}=10W/m^{2}'])

        figure; % plot time transient for output FWHM
        plot(t_vector,FWHM*1e9);
        title(['Output FWHM over time for peak ratio= ',
num2str(Ratio),', I_{2}=10W/m^{2}'])
        xlabel('time (s)')
        ylabel('Full-Width-Half Maximum (nm)')
```

```
        figure; % plot time transient for I1 peak intensity efficiency
        plot(t_vector,Eff_peak);
        title(['Peak Intensity Transmission in I_{1} over time for peak
ratio= ',num2str(Ratio),', I_{2}=10W/m^{2}'])
        xlabel('time (s)')
        ylabel('Peak Intensity Transmission')
    end

    FWHM_R(cnt_ratio)=FWHM(cnt_t+1);
% steady-state FWHM at certain ratio
    Eff_peak_R(cnt_ratio)=Eff_peak(cnt_t+1);
% steady-state I1 peak intensity at certain ratio
%    toc
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% plot trend w.r.t. Ratio %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
plot_trend=1; % set to 1 to plot trend
%%%%%%%%%%%%%%%%%%
if plot_trend == 1
    Peak_ratio= Ratio*I_scale2/I_scale1;
%%%%%%%% Full Width Half Maximum %%%%%%%
    figure; semilogx(Ratio,FWHM_R*1e9,'b');
    title(['FWHM over incident I_{2}/I_{1} ratio of SPP-NA=',
num2str(NA),' for Azo AML'])
    xlabel('I_{2}/I_{1} incident intensity ratio (fix I_{1})')
    ylabel('Full-Width-Half Maximum (nm)')

%    figure; semilogx(Peak_ratio,FWHM_R*1e9);
```

```
%    title('FWHM over focal-plane I_{2}/I_{1} ratio for Azo AML')
%    xlabel('I_{2}/I_{1} peak intensity ratio (fix I_{2})')
%    ylabel('Full-Width-Half Maximum (nm)')

%%%%%%% Efficiency %%%%%%%
    figure; semilogx(Ratio,Eff_peak_R,'b');
    title(['Peak Intensity Transmission in I_{1} over incident
I_{2}/I_{1} ratio of SPP-NA=',num2str(NA),' for Azo AML'])
    xlabel('I_{2}/I_{1} incident intensity ratio (fix I_{1})')
    ylabel('Peak Intensity Transmission')

%    figure; semilogx(Peak_ratio,Eff_peak_R);
%    title('Peak Intensity Transmission in I_{1} over focal-plane
I_{2}/I_{1} ratio for Azo AML')
%    xlabel('I_{2}/I_{1} peak intensity ratio (fix I_{2})')
%    ylabel('Peak Intensity Transmission')
end

%%%%%%%%%%%%%%%%%%%
%%%%% Save PSF %%%%%
%%%%%%%%%%%%%%%%%%%
%%% Note that the below are for only the last member of the vector: Ratio
save_PSF_file = 0; % set to 1 to save PSF values to a file.
%%%%%%%%%%%%%%%%%%%
if save_PSF_file == 1
    fid = fopen(filename1,'w');
    fprintf(fid, '%g %g %g %g\n', [r*1e9; I10/max(I10); I20/max(I20);
I1_matrix(length(z_vector),:)/max(I1_matrix(length(z_vector),:))]);
    fclose(fid);
    disp('File saved as:');
```

```
    disp(filename1);
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%% Save Ratio trend %%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
save_Ratio_file = 1; % set to 1 to save PSF values to a file.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if save_Ratio_file == 1
    fid = fopen(filename2,'w');
    fprintf(fid,'%s %s %s %s',['Incident_Ratio ','Focal_Plane_Ratio ',
'FWHM(nm) ','Efficiency']);
    fprintf(fid, '\n'); %%% TO FILE %%%
    fprintf(fid, '%g %g %g %g\n', [Ratio; Peak_ratio; FWHM_R*1e9;
Eff_peak_R]);
    fclose(fid);
    disp('File saved as:');
    disp(filename2);
end

if make_movie==1
    aviobj=close(aviobj);
end
```

A.2.2 The Main Function for Convergence Plots with Δz

```
%%%% AML_FD2 - convergence plot with dz
%%%% Main function to solve coupled PDE with finite difference
%%%% to plot convergence plots for L_1, L_2, L_\infty with dx
%%% May, 2010 (Sidney)

%close all;
clear all;

%%%% constants
Na = 6.022e23; h = 6.626e-34; c = 3e8;
l1 = 400e-9;
l2 = 532e-9;
moles1 = Na*h*c/l1; % energy of one mole of photons at lambda2
moles2 = Na*h*c/l2; % energy of one mole of photons at lambda2
Em1 = Na*h*c/l1; % energy of one mole of photons at lambda1
Em2 = Na*h*c/l2; % energy of one mole of photons at lambda2

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%   Parameters to adjust   %%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Ratio = 2000; % ratio of max(I2)/max(I1)
dr = 10e-9;  % r grid step size (unit = m)
I_in1 = 10; % W/m^2 at lambda1 (uniform intensity on dichromat)
I_in2 = 10; % W/m^2 at lambda2 % 10 W/m^2 is about correct
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% read from appropriate dichromat PSF file
data = load('PSFs_80.txt');
%%% Choose dichromat radius %%%
```

```
%R_D = 60e-6; %Radius of 0.83NA (160 zones) dichromat, ~60microns.
R_D = 39.2e-6;%Radius of 0.7NA (80 zones) dichromat, ~40microns.
%R_D = 26.5e-6;%Radius of 0.55NA (40 zones) dichromat,~26.5microns.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
r_t = data(1:20,1)*l1; I10_t = data(1:20,2); I20_t=data(1:20, 3);
clear data;
I10_t = I10_t/max(I10_t); I20_t = I20_t/max(I20_t);
% Calculate focal spot intensity scaling factor
% based on incident intensity
I_scale1 = I_in1*pi*R_D^2/energy_PSF(r_t,I10_t,0,max(r_t));
% Sidney:change moles1 scaling to coefficient, so I unit is still W/m^2
I10_t = I_scale1*I10_t;
I_scale2 = I_in2*pi*R_D^2/energy_PSF(r_t,I20_t,0,max(r_t));
I20_t = I_scale2*I20_t;
r = 0:dr:max(r_t);
I100 = interp1(r_t,I10_t,r,'cubic');
I200 = interp1(r_t,I20_t,r,'cubic');

%%%%%%%%%%%%%%
photochrome = 1;
%%%%%%%%%%%%%%
%%% Photochromic parameters
if photochrome == 1 % azobenzene polymer
    e1a=30000*log(10);
    e1b=12500*log(10);
    e2a=1250*log(10);
    e2b=3000*log(10);
    A0 =400.024; % mol/m^3
    Q1ab=.0037;
    Q1ba=.0021;
```

```
    Q2ab=.0021;
    Q2ba=.0072;
    Kba=.002;

elseif photochrome == 2 % thienylethene at 325nm & 633nm
    e1a=32675*log(10);
    e1b=11004*log(10);
    e2a=157.81*log(10);
    e2b=18133*log(10);
    A0 =306.044*2/2;
% mol/m^3 for initial absorbance of 2 with 200nm thickness
    Q1ab=0.2;
    Q1ba=9.6e-4;
    Q2ab=Q1ab;
    Q2ba=Q1ba;
    Kba=0;
end

alpha1A=e1a*Q1ab/Em1;
alpha2A=e2a*Q2ab/Em2;
alpha1B=e1b*Q1ba/Em1;
alpha2B=e2b*Q1ba/Em2;

% Defines the dz range for convergence plot
% Initializes infinity norm error vetors for I1, I2, and A
maxZg=8; %maximun gird size (2^)for z step
Linf_1=zeros(1,maxZg); %infinity error
Linf_2=zeros(1,maxZg);
Linf_A=zeros(1,maxZg);
```

```
zmax = 200e-9; % thickness of AML layer unit = meter
Zsample_max=floor(zmax/1/2^(maxZg-1)/1e-9);

for Zgrid=1:maxZg
%%% Choose step sizes and simulation ranges %%%
    dz = 1*2^(Zgrid-1)*1e-9; % 5nm step in z (unit = m)
    z_vector = 0:dz:zmax;
    Zsample_step= 2^(maxZg-Zgrid);
    Zsample=[Zsample_step:Zsample_step:Zsample_step*Zsample_max];
    dt = 0.1e-3; % 0.1 ms step in time t (unit = sec)
    tmax = 100e-3; % total simulation time, unit = s
    t_vector = 0:dt:tmax;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Create empty vectors
    I1_matrix = zeros(length(z_vector), length(r));
    I2_matrix = zeros(size(I1_matrix));
    A_matrix = A0*ones(length(z_vector), length(r));
    A_ctr=zeros(1,length(t_vector));
    A_ctr(1)=A_matrix(length(z_vector),1,1);
    FWHM = zeros(length(t_vector));
    half_FWHM_in = interp1(I100/max(I100),r,0.5,'cubic');
    Eff_peak = zeros(size(Ratio));
%% Peak (r=0) intensity efficiency of PSF
    Peak_in = I100(1); %% Incident peak intensity (lambda 1)
    I10=I100;
    I20=I200*Ratio;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    I1_matrix(1,:)=I10;
    I2_matrix(1,:)=I20;
    I1_matrix_s(1,:)=I10;
```

```
    I2_matrix_s(1,:)=I20;

    for cnt_z = 1:length(z_vector)-1
    %%% Initial condition for entire AML volume
I1_matrix(cnt_z+1,:)=(1-dz/2*e1a*A0)/(1+dz/2*e1a*A0)*I1_matrix(cnt_z,:);
I2_matrix(cnt_z+1,:)=(1-dz/2*e2a*A0)/(1+dz/2*e2a*A0)*I2_matrix(cnt_z,:);
    end
    FWHM(1) = 2*interp1(I1_matrix(length(z_vector),:)
/max(I1_matrix(length(z_vector),:)),r,0.5,'cubic');


    for cnt_t = 1:length(t_vector)-1 % step forward A to next time step
        A_prev=A_matrix;
        A_matrix= (1-dt/2*(I1_matrix*(alpha1A+alpha1B)+
I2_matrix*(alpha2A+alpha2B)+Kba))./(1+dt/2*(I1_matrix*(alpha1A+alpha1B)
+I2_matrix*(alpha2A+alpha2B)+Kba)).*A_prev
+dt*A0.*(I1_matrix*alpha1B+I2_matrix*alpha2B+Kba)./
(1+dt/2*(I1_matrix*(alpha1A+alpha1B)+I2_matrix*(alpha2A+alpha2B)+Kba));
        A_ctr(cnt_t+1)=A_matrix(length(z_vector),1);

        for cnt_z = 1:length(z_vector)-1
        %%% Initial condition for entire AML volume
I1_matrix(cnt_z+1,:)=(1-dz/2.*((e1a-e1b).*A_matrix(cnt_z,:)+e1b*A0))./
(1+dz/2.*((e1a-e1b).*A_matrix(cnt_z,:)+e1b*A0)).*I1_matrix(cnt_z,:);
I2_matrix(cnt_z+1,:)=(1-dz/2.*((e2a-e2b).*A_matrix(cnt_z,:)+e2b*A0))./
(1+dz/2.*((e2a-e2b).*A_matrix(cnt_z,:)+e2b*A0)).*I2_matrix(cnt_z,:);
        end
        FWHM(cnt_t+1) = 2*interp1(I1_matrix(length(z_vector),:)
/max(I1_matrix(length(z_vector),:)),r,0.5,'cubic');
        %FWHM at AML output (z=zmax)
```

```
        Eff_peak(cnt_t+1) = I1_matrix(length(z_vector),1)/Peak_in;
        %Peak intensity efficiency at AML output (z=zmax)
    end

    if Zgrid==1
    % use smallest step solution as exact solution
        I_1exact=I1_matrix(Zsample,:);
        I_2exact=I2_matrix(Zsample,:);
        A_exact=A_matrix(Zsample,:);
    else
        Z_plot(Zgrid)=dz; %record dz vector
        % infinity norm for error vectors
        Linf_1(Zgrid)=max(max(abs(I1_matrix(Zsample,:)-I_1exact)));
        Linf_2(Zgrid)=max(max(abs(I2_matrix(Zsample,:)-I_2exact)));
        Linf_A(Zgrid)=max(max(abs(A_matrix(Zsample,:)-A_exact)));
        % L2 norm for error vectors
        L2_1(Zgrid)=sqrt(sum(sum((I1_matrix(Zsample,:)-I_1exact).*
(I1_matrix(Zsample,:)-I_1exact)))/size(I1_matrix,1)/size(I1_matrix,2));
        L2_2(Zgrid)=sqrt(sum(sum((I2_matrix(Zsample,:)-I_2exact).*
(I2_matrix(Zsample,:)-I_2exact)))/size(I2_matrix,1)/size(I2_matrix,2));
        L2_A(Zgrid)=sqrt(sum(sum((A_matrix(Zsample,:)-A_exact).*
(A_matrix(Zsample,:)-A_exact)))/size(A_matrix,1)/size(A_matrix,2));
        % L1 norm for error vectors
        L1_1(Zgrid)=sum(sum(abs(I1_matrix(Zsample,:)-I_1exact)))
/size(I1_matrix,1)/size(I1_matrix,2);
        L1_2(Zgrid)=sum(sum(abs(I2_matrix(Zsample,:)-I_2exact)))
/size(I2_matrix,1)/size(I2_matrix,2);
        L1_A(Zgrid)=sum(sum(abs(A_matrix(Zsample,:)-A_exact)))
/size(A_matrix,1)/size(A_matrix,2);
    end
```

```
end

Peak_ratio= Ratio*I_scale2/I_scale1;
% I_scale2/I_scale1
% to see how much peak ratio is reduced from incident
% ratio =0.2512 for N80

%%%%%%%%%%%%%%%%%%%%%%
%%% plot_converge=1; %%%
%%%%%%%%%%%%%%%%%%%%%%
if plot_converge==1
    figure;
    loglog(Z_plot*1e9,Linf_1/1e4,'xb',Z_plot*1e9,Linf_2/1e6,'xg',
Z_plot*1e9,Linf_A,'xr',Z_plot*1e9,Z_plot*1e9,'k',
Z_plot*1e9,Z_plot.^2*5e17,'k--','LineWidth',2)
    title(['Global Convergence Plot L_\infty with dz (zmax=200nm),
Ratio=',num2str(Ratio),', dt=',num2str(dt),', tmax=',num2str(tmax)])
    legend('L_\infty/1e4 for I_1','L_\infty/1e6 for I_2',
'L_\infty for A','first order','second order')
    xlabel('dz(nm)'); ylabel('error')
    figure;
    loglog(Z_plot*1e9,L2_1/1e4,'xb',Z_plot*1e9,L2_2/1e6,'xg',Z_plot*1e9,
L2_A,'xr',Z_plot*1e9,Z_plot*1e9,'k',Z_plot*1e9,Z_plot.^2*5e16,
'k--','LineWidth',2)
    title(['Global Convergence Plot L_2 with dz (zmax=200nm),
Ratio=',num2str(Ratio),', dt=',num2str(dt),', tmax=',num2str(tmax)])
    legend('L_2/1e4 for I_1','L_2/1e6 for I_2','L_2 for A',
'first order','second order')
    xlabel('dz(nm)'); ylabel('error')
    figure;
```

```
    loglog(Z_plot*1e9,L1_1/1e4,'xb',Z_plot*1e9,L1_2/1e6,'xg',Z_plot*1e9,
L1_A,'xr',Z_plot*1e9,Z_plot*1e8,'k',Z_plot*1e9,Z_plot.^2*5e14,
'k--','LineWidth',2)
    title(['Global Convergence Plot L_1 with dz (zmax=200nm),
Ratio=',num2str(Ratio),', dt=',num2str(dt),', tmax=',num2str(tmax)])
    legend('L_1/1e4 for I_1','L_1/1e6 for I_2','L_1 for A',
'first order','second order')
    xlabel('dz(nm)'); ylabel('error')
end

%%% Plot intensity profile crossections when plot_cross ==1 %%%
plot_cross =0;
%%%%%%%%%%%%%
if plot_cross ==1
    figure;
    subplot(3,1,1);
    imagesc(r*1e9, z_vector*1e9, I1_matrix); title('I_1'); colorbar;
    xlabel('r(nm)'); ylabel('z(nm)');
    axis([0,750,0,200])
    subplot(3,1,2);
    imagesc(r*1e9, z_vector*1e9, I2_matrix); title('I_2'); colorbar;
    xlabel('r(nm)'); ylabel('z(nm)');
    axis([0,750,0,200])

    A_matrix = A0*(Q1ba*e1b*I1_matrix + Q2ba*e2b*I2_matrix + Kba)./
((Q1ab*e1a+Q1ba*e1b)*I1_matrix + (Q2ab*e2a+Q2ba*e2b)*I2_matrix+Kba);
    subplot(3,1,3);
    imagesc(r*1e9, z_vector*1e9, A_matrix); title('[A]'); colorbar;
    xlabel('r(nm)'); ylabel('z(nm)');
    axis([0,750,0,200])
```

```
    figure;
    plot(r*1e9, I10/max(I10), r*1e9, I20/max(I20), r*1e9,
    I1_matrix(length(z_vector),:)/max(I1_matrix(length(z_vector),:)));
    title(['Azo-AML=200nm,incident I_{2}=10W/m^{2},I_{2}/I_{1}
    Ratio= ',num2str(Ratio)])
    legend('I_{1}','I_{2}','I_{out}')

    for cnt_z = 1:length(z_vector)
        FWHM_cross(cnt_z) = 2*interp1(I1_matrix(cnt_z,:)
/max(I1_matrix(cnt_z,:)),r,0.5,'cubic');
    end
    figure;
    plot(z_vector*1e9,FWHM_cross*1e9);
    title(['FWHM over AML thickness for peak ratio= ',num2str(Peak_ratio)])
    xlabel('z(nm)')
    ylabel('Full-Width-Half Maximum (nm)')
end

%%% Plot transient w.r.t. time when plot_transient ==1 %%%
plot_transient =0;
%%%%%%%%%%%%%%%%%
if plot_transient ==1
    figure;
    plot(t_vector,A_ctr)
    xlabel('time (s)')
    ylabel('A matrix')
    title(['Central (r=0) absorbance in AML over time for peak ratio= ',
num2str(Peak_ratio)])
    figure; plot(t_vector,FWHM*1e9);
```

```
    title(['Output FWHM over time for peak ratio= ',num2str(Peak_ratio)])
    xlabel('time (s)')
    ylabel('Full-Width-Half Maximum (nm)')

    figure; plot(t_vector,Eff_peak);
    title(['Peak Intensity Transmission in I_{1} over time for
peak ratio= ',num2str(Peak_ratio)])
    xlabel('time (s)')
    ylabel('Peak Intensity Transmission')
end

save_values_file = 0; % set to 1 to save PSF values to a file.
if save_values_file == 1
%%%% enter appropriate filenmae here to save PSF values
    filename2 = '20100509PSFs_80_FWHM_Eff.txt';
    fid = fopen(filename2,'w');
    fprintf(fid,'%s %s %s %s',
['Incident_Ratio ','Focal_Plane_Ratio ','FWHM(nm) ','Efficiency']);
    fprintf(fid, '\n'); %%% TO FILE %%%
    fprintf(fid, '%g %g %g %g\n', [Ratio; Peak_ratio; FWHM*1e9; Eff_peak]);
    fclose(fid);
    disp('File saved as:');
    disp(filename2);
end
```

A.2.3 The Main Function for Convergence Plots with Δt

```
%%% AML_FD3 - convergence plot with dt
%%%% Main function to solve coupled PDE with finite difference
%%%% to plot convergence plots for L_1, L_2, L_\infty with dt
%%% May, 2010 (Sidney)

close all;
clear all;

%%%% constants
Na = 6.022e23; h = 6.626e-34; c = 3e8;
l1 = 400e-9;
l2 = 532e-9;
moles1 = Na*h*c/l1; % energy of one mole of photons at lambda2
moles2 = Na*h*c/l2; % energy of one mole of photons at lambda2
Em1 = Na*h*c/l1; % energy of one mole of photons at lambda2
Em2 = Na*h*c/l2; % energy of one mole of photons at lambda2

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%   Parameters to adjust   %%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Ratio = 2000; % ratio of max(I2)/max(I1)
dr = 10e-9;  % r grid step size (unit = m)
I_in1 = 10; % W/m^2 at lambda1 (uniform intensity on dichromat)
I_in2 = 10; % W/m^2 at lambda2 % 10 W/m^2 is about correct
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% read from appropriate dichromat PSF file
data = load('PSFs_80.txt');
%%% Choose dichromat radius %%%
```

```
%R_D = 60e-6; %Radius of 0.83NA (160 zones) dichromat, ~60microns.
R_D = 39.2e-6;%Radius of 0.7NA (80 zones) dichromat, ~40microns.
%R_D = 26.5e-6;%Radius of 0.55NA (40 zones) dichromat,~26.5microns.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
r_t = data(1:20,1)*l1; I10_t = data(1:20,2); I20_t=data(1:20, 3);
clear data;
I10_t = I10_t/max(I10_t); I20_t = I20_t/max(I20_t);
% Calculate focal spot intensity scaling factor
% based on incident intensity
I_scale1 = I_in1*pi*R_D^2/energy_PSF(r_t,I10_t,0,max(r_t));
% Sidney:change moles1 scaling to coefficient, so I unit is still W/m^2
I10_t = I_scale1*I10_t;
I_scale2 = I_in2*pi*R_D^2/energy_PSF(r_t,I20_t,0,max(r_t));
I20_t = I_scale2*I20_t;
r = 0:dr:max(r_t);
I100 = interp1(r_t,I10_t,r,'cubic');
I200 = interp1(r_t,I20_t,r,'cubic');

%%%%%%%%%%%%%%
photochrome = 1;
%%%%%%%%%%%%%%
%%% Photochromic parameters
if photochrome == 1 % azobenzene polymer
    e1a=30000*log(10);
    e1b=12500*log(10);
    e2a=1250*log(10);
    e2b=3000*log(10);
    A0 =400.024; % mol/m^3
    Q1ab=.0037;
    Q1ba=.0021;
```

```
    Q2ab=.0021;
    Q2ba=.0072;
    Kba=.002;

elseif photochrome == 2 % thienylethene at 325nm & 633nm
    e1a=32675*log(10);
    e1b=11004*log(10);
    e2a=157.81*log(10);
    e2b=18133*log(10);
    A0 =306.044*2/2; % mol/m^3 for initial absorbance of 2 with 200nm thickness
    Q1ab=0.2;
    Q1ba=9.6e-4;
    Q2ab=Q1ab;
    Q2ba=Q1ba;
    Kba=0;
end

alpha1A=e1a*Q1ab/Em1;
alpha2A=e2a*Q2ab/Em2;
alpha1B=e1b*Q1ba/Em1;
alpha2B=e2b*Q1ba/Em2;

maxTg=12; %maximun gird size (2^)for t step
Linf_1=zeros(1,maxTg); %infinity error
Linf_2=zeros(1,maxTg);
Linf_A=zeros(1,maxTg);

zmax = 200e-9; % thickness of AML layer unit = meter

for Tgrid=1:maxTg
```

```
%%% Choose step sizes and simulation ranges %%%
    dz = 5e-9; % 5nm step in z (unit = m)
    z_vector = 0:dz:zmax;
    %dt = 2^(Tgrid-1)*0.001e-3; % 0.1 ms step in time t (unit = sec)
    tmax = 100e-3; % total simulation time, unit = s
    dt=tmax/10/2^(maxTg-Tgrid);
    t_vector = 0:dt:tmax;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Create empty vectors %%%
    I1_matrix = zeros(length(z_vector), length(r));
    I2_matrix = zeros(size(I1_matrix));
    A_matrix = A0*ones(length(z_vector), length(r));
    A_ctr(1)=A_matrix(length(z_vector),1,1);
    half_FWHM_in = interp1(I100/max(I100),r,0.5,'cubic'); %% Sidney
    Peak_in = I100(1); %% Incident peak intensity (lambda 1)
    I10=I100;
    I20=I200*Ratio;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    I1_matrix(1,:)=I10;
    I2_matrix(1,:)=I20;
    I1_matrix_s(1,:)=I10;
    I2_matrix_s(1,:)=I20;

    for cnt_z = 1:length(z_vector)-1
%%% Initial condition for entire AML volume
I1_matrix(cnt_z+1,:)=(1-dz/2*e1a*A0)/(1+dz/2*e1a*A0)*I1_matrix(cnt_z,:);
I2_matrix(cnt_z+1,:)=(1-dz/2*e2a*A0)/(1+dz/2*e2a*A0)*I2_matrix(cnt_z,:);
    end
    FWHM(1) = 2*interp1(I1_matrix(length(z_vector),:)
/max(I1_matrix(length(z_vector),:)),r,0.5,'cubic');
```

```
    for cnt_t = 1:length(t_vector)-1
        A_prev=A_matrix;
        A_matrix= (1-dt/2*(I1_matrix*(alpha1A+alpha1B)+
I2_matrix*(alpha2A+alpha2B)+Kba))./(1+dt/2*(I1_matrix*(alpha1A+alpha1B)
+I2_matrix*(alpha2A+alpha2B)+Kba)).*A_prev
+dt*A0.*(I1_matrix*alpha1B+I2_matrix*alpha2B+Kba)./
(1+dt/2*(I1_matrix*(alpha1A+alpha1B)+I2_matrix*(alpha2A+alpha2B)+Kba));

        for cnt_z = 1:length(z_vector)-1
        %%% Initial condition for entire AML volume
I1_matrix(cnt_z+1,:)=(1-dz/2.*((e1a-e1b).*A_matrix(cnt_z,:)+e1b*A0))./
(1+dz/2.*((e1a-e1b).*A_matrix(cnt_z,:)+e1b*A0)).*I1_matrix(cnt_z,:);
I2_matrix(cnt_z+1,:)=(1-dz/2.*((e2a-e2b).*A_matrix(cnt_z,:)+e2b*A0))./
(1+dz/2.*((e2a-e2b).*A_matrix(cnt_z,:)+e2b*A0)).*I2_matrix(cnt_z,:);
        end
    end

    if Tgrid==1
    % use smallest step solution as exact
        I_1exact=I1_matrix;
        I_2exact=I2_matrix;
        A_exact=A_matrix;
    else
        T_plot(Tgrid)=dt; %record dt vector
        % infinity norm for error vectors
        Linf_1(Tgrid)=max(max(abs(I1_matrix-I_1exact)));
        Linf_2(Tgrid)=max(max(abs(I2_matrix-I_2exact)));
        Linf_A(Tgrid)=max(max(abs(A_matrix-A_exact)));
        % L2 norm for error vectors
```

```
        L2_1(Tgrid)=sqrt(sum(sum((I1_matrix-I_1exact).*
(I1_matrix-I_1exact)))/size(I1_matrix,1)/size(I1_matrix,2));
        L2_2(Tgrid)=sqrt(sum(sum((I2_matrix-I_2exact).*
(I2_matrix-I_2exact)))/size(I2_matrix,1)/size(I2_matrix,2));
        L2_A(Tgrid)=sqrt(sum(sum((A_matrix-A_exact).*
(A_matrix-A_exact)))/size(A_matrix,1)/size(A_matrix,2));
        % L1 norm for error vectors
        L1_1(Tgrid)=sum(sum(abs(I1_matrix-I_1exact)))
/size(I1_matrix,1)/size(I1_matrix,2);
        L1_2(Tgrid)=sum(sum(abs(I2_matrix-I_2exact)))
/size(I2_matrix,1)/size(I2_matrix,2);
        L1_A(Tgrid)=sum(sum(abs(A_matrix-A_exact)))
/size(A_matrix,1)/size(A_matrix,2);
    end
end

%%%%%%%%%%%%%%%%%%%%%%
%%% plot_converge=1; %%%
%%%%%%%%%%%%%%%%%%%%%%
if plot_converge==1
    figure;
    loglog(T_plot,Linf_1/1e2,'xb',T_plot,Linf_2/1e4,'xg',
T_plot,Linf_A,'xr',T_plot,T_plot,'k','LineWidth',2)
    title(['Global Convergence Plot L_\infty with dz (zmax=200nm),
Ratio=',num2str(Ratio),', dz=',num2str(dz),'nm, tmax=',num2str(tmax)])
    legend('L_\infty/1e2 for I_1','L_\infty/1e4 for I_2',
'L_\infty for A','first order')
    xlabel('dt(sec)'); ylabel('error')
    figure;
loglog(T_plot,L2_1/1e2,'xb',T_plot,L2_2/1e4,'xg',T_plot,L2_A,'xr',
```

```
T_plot,T_plot,'k','LineWidth',2)
    title(['Global Convergence Plot L_2 with dz (zmax=200nm),
Ratio=',num2str(Ratio),', dz=',num2str(dz),'nm, tmax=',num2str(tmax)])
    legend('L_2/1e2 for I_1','L_2/1e4 for I_2','L_2 for A','1st order')
    xlabel('dt(sec)'); ylabel('error')
    figure;
    loglog(T_plot,L1_1/1e2,'xb',T_plot,L1_2/1e4,'xg',
T_plot,L1_A,'xr',T_plot,T_plot,'k','LineWidth',2)
    title(['Global Convergence Plot L_1 with dz (zmax=200nm),
Ratio=',num2str(Ratio),', dz=',num2str(dz),'nm, tmax=',num2str(tmax)])
    legend('L_1/1e2 for I_1','L_1/1e4 for I_2','L_1 for A','1st order')
    xlabel('dt(sec)'); ylabel('error')
end
```

A.2.4 Point-Spread-Function Energy Calculation

```
function y  = energy_PSF(r,I,a,b)
%%% integrate I*r*dr from a to b.
    y = quadv(@PSF_integrand,a,b);
    function w = PSF_integrand(x)
    %%% interpolating the integrand.
       w = interp1(r, I, x, 'cubic').*x;
    end
end
```

Appendix B

Matlab Genetic Algorithm Implimentation Code and Results

B.1 Main Function for Genetic Algorithm

```
%%%% Main GA code and initialize parameters
%%%% 08/08/2007 by Paul Rogge
%%%%% the optimization variables are the radii. Fix phase height.
%%% This makes fabrication of multiple NA dichromats easier.
%%%% 12/26/2007 by Rajesh Menon
%%% N=80 run time ~11hr
%%% N=160 run time ~21.5hr

function x = gen_alg_dual_wavelength2(Nzones)
tic
% Define parameters
l2 = 532/405;
%Normalize lambda2 wrt lambda1 (l2/l1) Sidney changed l1 to 405 from 400
l3 = 633/405; % normalize lambda3 wrt lambda1
N = Nzones;
z = 100;
```

```
Rn = sqrt(Nzones*z + (Nzones/2)^2);
NA = sin(atan(sqrt(Nzones*z + (Nzones/2).^2)/z));

%n_refr1 = 1.4702;
%n_refr2 = 1.4608;
n_refr1 = 1.4205;
% refractive index for l1=405nm, HSQ 450C 2hr furnace bake
n_refr2 = 1.4100;
% refrative index for l2=532nm, HSQ 450C 2hr furnace bake
%n_refr1 = 1.501372;
% refrative index for l1 = 400 nm, measured for PMMA in anisole
%n_refr2 = 1.487042;
% refrative index for l2 = 532 nm, measured for PMMA in anisole
%%% measurements done with ISN spectroscopic ellipsometer by Will Arora

%%% phase height for pi phase shift at lambda1
h_pi = 1/(2*(real(n_refr1)-1))
%%% rho1, rho2, rho3 for spot and doughnut
s1 = .5/NA;
s2 = 0.5*l2/NA;
s3 = 1.22*l2/NA-s2;
%%% population size
pop_size = 25;

% fk is generated as follows for fft
Xmax = 2*Rn;
dx = .25;
dx = Xmax/(round(Xmax/dx));
xvec = -Xmax:dx:Xmax;
```

```
[xx, yy] =meshgrid(xvec, xvec);
dd = sqrt(xx.^2 + yy.^2);
d = sqrt(xx.^2 + yy.^2 + z^2);
H = exp(i*2*pi*d)./d.*(1 + z./d);
fk = fft2(H);

% lambda 2 parameters
dd2 = sqrt((xx/l2).^2 + (yy/l2).^2);
d2 = sqrt((xx/l2).^2 + (yy/l2).^2 + (z/l2)^2);
H2 = exp(i*2*pi*d2)./d2.*(1 + (z/l2)./d2);
fk2 = fft2(H2);

% Set up Genetic Algorithm (GA)
% Define fabrication constraint
delta = .5;
% Define the inequality contraints: A*x <= b
% Define the A matrix (Nx1): [-r(n-1) + r(n)]
A = zeros([N, N]);
for m = 1:N-1
    A(m,m) = 1;
    A(m,m+1) = -1;
end
A(N,N) = 1;
% Define b matrix(Nx1): [delta ... UB(r) UB(h)]
b = -delta.*(ones([N, 1]));
b(N,1) = Rn; % upper bound for r
% Define lower bounds
LB = zeros(1,N);
LB(1) = 2*delta; % first radius must be > 2*delta
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
% Keep track of best X values
best_X= zeros(1,Nzones);
best_E = 3.4e+35; % some large number, so it gets updated the first time around
best_E2E1=1;
best_Enod=1e20;
fid = fopen('progress-12-07-10.txt','a'); %%% TO FILE %%%
fgood=fopen('improve-12-07-10.txt','a'); %%% TO FILE %%%

fprintf(fid,'%s %s %s %s %s %s\n',
['GAcount ','Energy ','E1 ','E2 ','Enod ','x']);
fprintf(fid, '\n'); %%% TO FILE %%%
fprintf(fgood, '%s %s %s %s %s %s\n',
['Type ','GAcount ','Energy ','E1 ','E2 ','Enod']);
fprintf(fgood, '\n'); %%% TO FILE %%%

%%% Exmaple: use an HSQ Zone-Plate as initial population
% N=Nzones ZP in HSQ
z = 100; %40.5um focal length
n = 1:Nzones;
rn = sqrt(n*z + (n/2).^2);
init_pop0 = [rn 1.189];

%%% To run 50 rounds of GA %%%
for GAcount=1:50 %%% TO FILE %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

    if length(init_pop0) < Nzones
        %%% add sequentially after the last element to pad.
        init_pop = [init_pop0 ceil(max(init_pop0))
:(Nzones-1-length(init_pop0)+ceil(max(init_pop0)))];
```

```
    elseif length(init_pop0) > Nzones
        init_pop = init_pop0(1:Nzones);
    elseif length(init_pop0) == Nzones
        init_pop = init_pop0;
    end

    % create random pop based on initial vector
    n = ones(pop_size,1);
    init_pop = n*init_pop;
    mu = 0;
    sigma_r = 0.4*405/405;%%%% 405????
    delta_r = normrnd(mu, sigma_r, pop_size, N);
    % preserve initial population!
    % delta_r(1,1:Nzones)=zeros(1,Nzones);

    init_pop = init_pop + delta_r;

%%%%%%%%%%%%%%%%%%%%%%%%%
%%%% Define GA options  %%%%
%%%%%%%%%%%%%%%%%%%%%%%%%
    options = gaoptimset;
    %Set population size
    options = gaoptimset(options, 'PopulationSize', pop_size);
    % if no improvement in this time, quits
    options = gaoptimset(options, 'StallTimeLimit', Inf);
    options = gaoptimset(options, 'Display', 'iter'); % displays each iteration
    options = gaoptimset(options, 'MutationFcn', @mutationadaptfeasible);
    options = gaoptimset(options, 'InitialPopulation', init_pop);
    options = gaoptimset(options, 'Generations', 100);
    options = gaoptimset(options, 'CrossoverFraction', .4);
```

```
    options = gaoptimset(options, 'Vectorize','on');
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

    % fitness function
    FitnessFunction = @(x) energy_function_RM
    (x, h_pi, n_refr1, n_refr2, s1, s2, s3, l2, fk, fk2, xvec, dd, dd2);
    [x, fval, exitflag, output, final_pop] =
ga(FitnessFunction, N, A, b, [], [], LB, [], [], options);
    [Energy,E1,E2,Enod]=energy_function_RM
(x, h_pi, n_refr1, n_refr2, s1, s2, s3, l2, fk, fk2, xvec, dd, dd2)

    fprintf(fid, '%g %g %g %g %g ', [GAcount,Energy,E1,E2,Enod]);
    fprintf(fid,'%g ',x);
    fprintf(fid, '\n');

%%%%%%%%%%%%% Compare %%%%%%%%%%%%% %%% TO FILE %%%
    if Energy < best_E
        fprintf(fgood, '%g %g %g %g %g %g', [1,GAcount,Energy,E1,E2,Enod]);
        fprintf(fgood, '\n');
        best_E = Energy;
    elseif E2/E1 > best_E2E1
        fprintf(fgood, '%g %g %g %g %g %g', [2,GAcount,Energy,E1,E2,Enod]);
        fprintf(fgood, '\n');
        best_E2E1 = E2/E1;
    elseif Enod < best_Enod
        fprintf(fgood, '%g %g %g %g %g %g', [3,GAcount,Energy,E1,E2,Enod]);
        fprintf(fgood, '\n');
        best_Enod = Enod;
    else
```

```
        fprintf(fgood, '%g %g', [4,GAcount]);
        fprintf(fgood, '\n');
    end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    GAcount
    best_E
    best_E2E1
    best_Enod
end
fclose(fid);
fclose(fgood);
toc
end
```

B.2 Fitness Function Energy Computation

```
% compute total energy of all vectors created by adding random noise
% to an initial vector created by the GA for dual lambda
% Paul Rogge
% 8/8/07
% Updated by Rajesh Menon October 19, 2007.
% modify, fk, fk2, z, dd, dd2 before calling this
%%% Modified 12/26/2007 such that h is not an optimization variable

function [E,energy_0_s1,energy_s2_s3,E120] = energy_function_RM
(x_noise, h, n_refr1, n_refr2, s1, s2, s3, l2, fk, fk2, xvec, dd, dd2)

% Compute energy for lambda 1 intensity funct.
% Returns vector of point spread function and vector rho=[0:dx:2*max(r)]
[PSF_matrix1, rho1] = intensity_diff_lens_speed_vectortest
```

```
(x_noise, h, n_refr1, fk, xvec, dd);

% Integrate to calculate energy of PSF from 0 to s1
energy_0_s1 = quadv(@PSF_integrand2, 0, s1);
%max_E_0_s1=max(max(energy_0_s1))% Sidney: print max energy for blue

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function w = PSF_integrand2(x)
    w = interp1(rho1, PSF_matrix1', x, 'linear').*x;
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Compute energy for lambda 2
% intensity funct. returns vector of point spread function and
% vector rho which is [0:dx:2*max(r)]

%Inputs must be normalized to lambda2
[PSF_matrix2, rho2] = intensity_diff_lens_speed_vectortest
(x_noise/l2, h/l2, n_refr2, fk2, xvec/l2, dd2);
rho2 = rho2*l2; % to keep spatial coord. in respect to lambda1

% Integrate to calculate energy of PSF from s2 to s3
energy_s2_s3 = quadv(@PSF_integrand3, s2, s3);
%max_E_s2_s3 = max(max(energy_s2_s3)) % Sidney: print max energy for green

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function z = PSF_integrand3(x)
    z = interp1(rho2, PSF_matrix2', x, 'linear').*x;
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
[ny, nx] = size(x_noise);
```

```
% find I(rho=0) for l1 and l2
El20 = PSF_matrix2(1:ny, 1)'; % lambda2 psf @ rho = 0
%min_El20=min(min(El20)) % Sidney: print min field for node

% compute energy of the original vector
% weight PSF(rho = 0) to replace nonlinear constraints
%%%%%%%%%%%%%%%%%%%%%%
%%% Set Weight Here %%%
%%%%%%%%%%%%%%%%%%%%%%
wwww = [1e5 1 10];
%%%%%%%%%%%%%%%%%%%%%%
% weights for optimization
% wwww(1) min node, wwww(2) max spot, wwww(3) max doughnut
E_original = wwww(1)*El20 - wwww(2)*energy_0_s1 - wwww(3)*energy_s2_s3;
E = E_original';
```

B.3 Transmission Function Computation

```
function T = transmission_diff_lens(dd, r, phi)
% Computes which zones are in phase or out of phase in a diffraction lens
% r is a vector containing the radii of the lens
% x & y must be vectors
% phi relates the zones of the lens that are out of phase

N = length(r);
const = exp(i*phi);
T = 0;
for k = 1:N-1
    rr = r(k);
```

```
    T = T + (-1)^(k-1)*(dd<=rr);
end
T = T*(1-const);
rr = r(N);
if rem(N,2) == 0
    T = T + const*(dd<=rr);
else
    T = T + (dd<=rr);
end
%%% code below to graph
% figure; imagesc(real(T)); axis equal; colorbar;
```

B.4 Focal-Plane Intensity Caculation

B.4.1 Main Function

```
%%%  Main code to drive the calulation of intensity near/at the
%%% focal plane of a diffractive lens
%%% Rajesh Menon 01/07/05
%%% updated 01/08/05 RM
%%% updated 03/19/07 RM
%%% updated 06/20/07  PR

function [PSF_matrix, rho] =
intensity_diff_lens_speed_vectortest(x, h, n_refr, fk, xvec, dd)

%%% Note all space coordinates are normalized to lambda
% x is matrix, pop size x length(x)
[matrix_y, N] = size(x);
```

```
% find rho, which is same for all PSF's
index = find(abs(xvec) == min(abs(xvec)));
%Find the center point (0,PSF(0))
rho = xvec(index:(length(xvec)-1)); %vector from center to end
% loop through each x vector to find its corresponding PSF.
PSF_matrix = zeros(matrix_y, N);

for row_number = 1:matrix_y
    % get x vector from matrix
    r = x(row_number, 1:N);
    % T calculates the transmission function of the lens
    phi = (real(n_refr)-1)*h*2*pi;
    % phi in terms of h, normalized to lambda1
    T = transmission_diff_lens(dd, r, phi);

    ZZP = fresnel_diffr4(T, fk); %Compute FFT
    I = abs(ZZP).^2; % Intensity function
    %I = I/max(max(I)); %% uncomment this if we dont want to normalize

% take slice of intensity distribution from center to end
% (circular symmetric)
    PSF = I(index+1,(index+1):(length(xvec)));
    % 1-D slice of point spread function
   % Store PSF in the PSF matrix
    PSF_matrix(row_number, 1:length(PSF)) = PSF;
end

    plot_figs = 0; % make 1 if you want the figures plotted.
if plot_figs == 1
    %figure; imagesc(xvec, xvec, real(T)); axis equal; colorbar;
```

```
    figure; imagesc(xvec, xvec, I/max(max(I))); axis equal; colorbar;
    colormap('gray'); axis([-5 5 -5 5]);
    %figure; plot(rho, PSF, 'g');
    end
end
```

B.4.2 FFT Function

```
function g=fresnel_diffr4(f,fk)
% fk is generated as follows
% [xx, yy] =meshgrid(xvec, yvec);
% d = sqrt(xx.^2+yy.^2 + z^2);
% H = exp(i*2*pi*d)./d.*(1+z./d);
% fk = fft2(H)
g = fftshift(fft2(fft2(f).*fk));
```

Appendix C

Labview Control Code for Absorbance Modulation Imaging

C.1 Front panel of Labview VI control

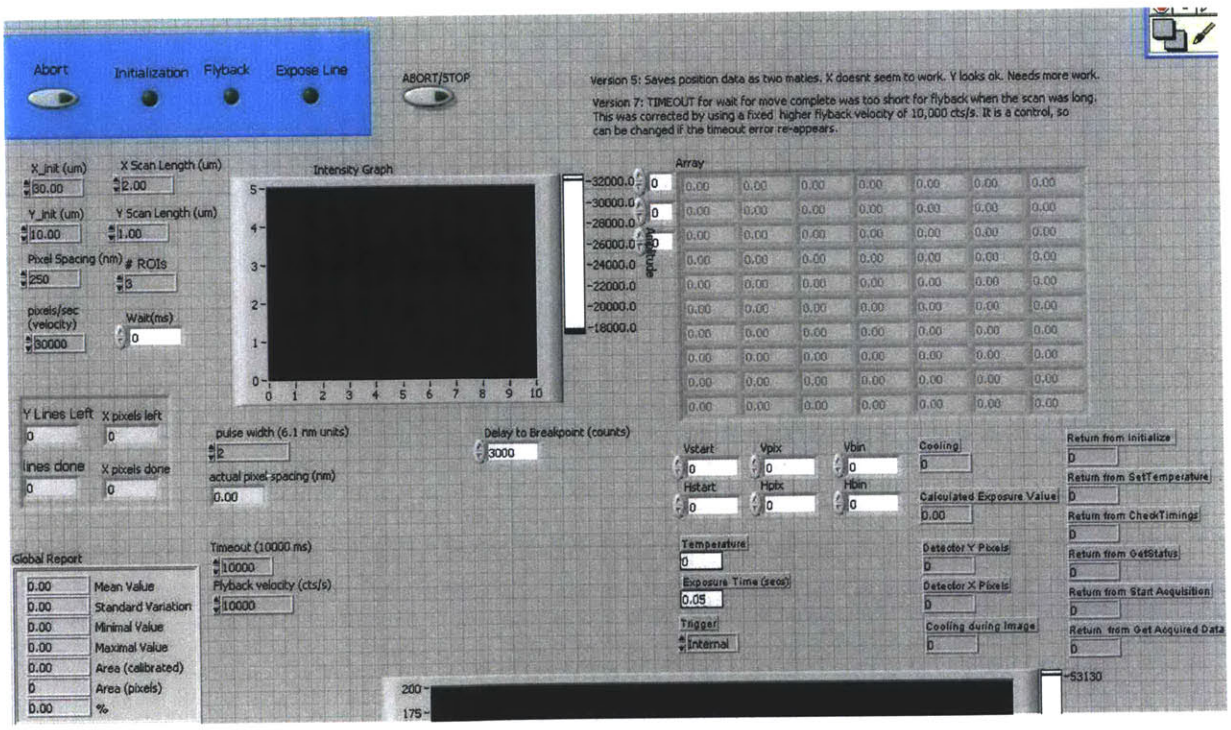

Figure C-1: Front panel of Labview VI control for camera image acquisition and stage scanning. The file name of the VI is "ZPAM 2D Sidney 2010.vi"

C.2 Initialization of Andor Clara CDD

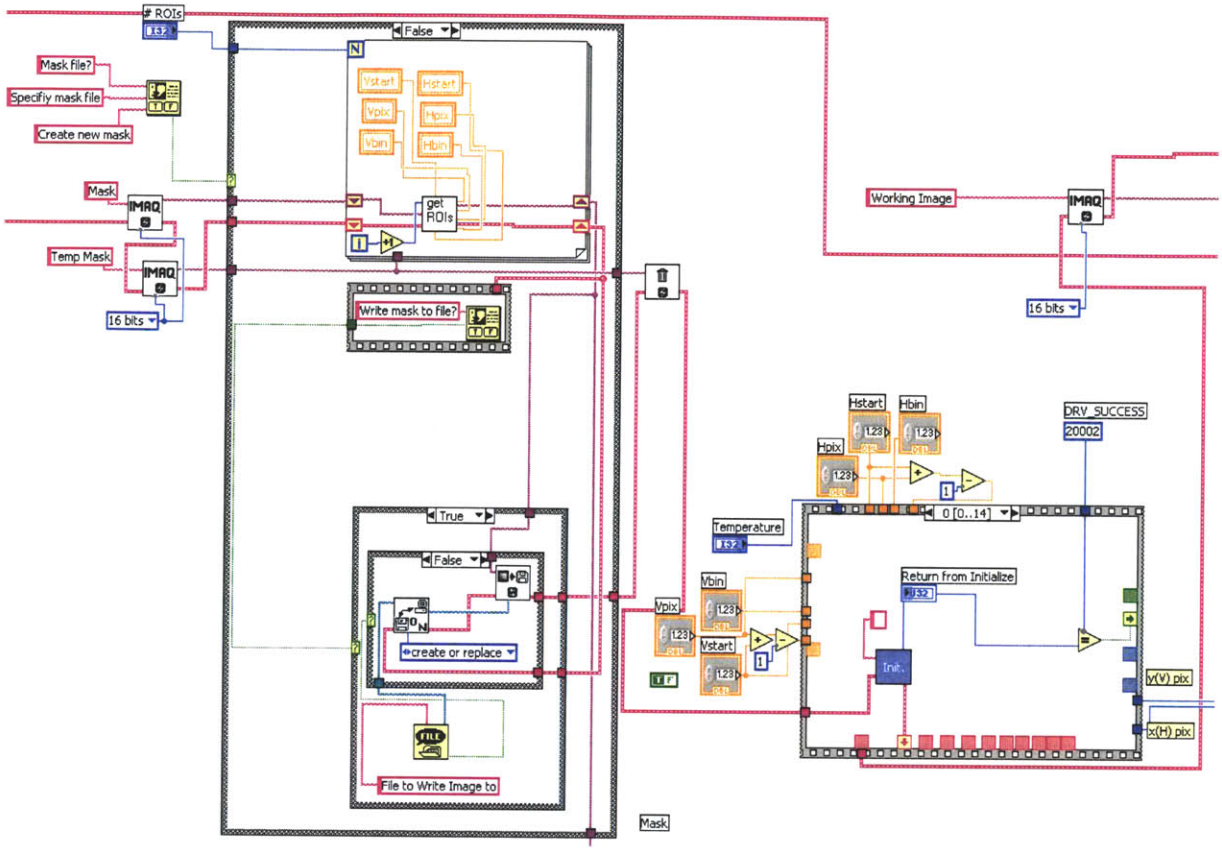

Figure C-2: Block diagram for region of interest (ROI) definition and Andor Clara CCD camera initialization

Figure C-3-C-17 shows the initialization frame which defines the region of interest (ROI) and initialize the CCD camera for following image acquisition. ROI is defined by either loading from existing files or taking a test frame with the CCD camera and drawing circles of interest by hand. The CCD camera undergoes the same initialization sequence as will be shown in the following steps. The CCD camera is turned off before exiting the ROI definition block.

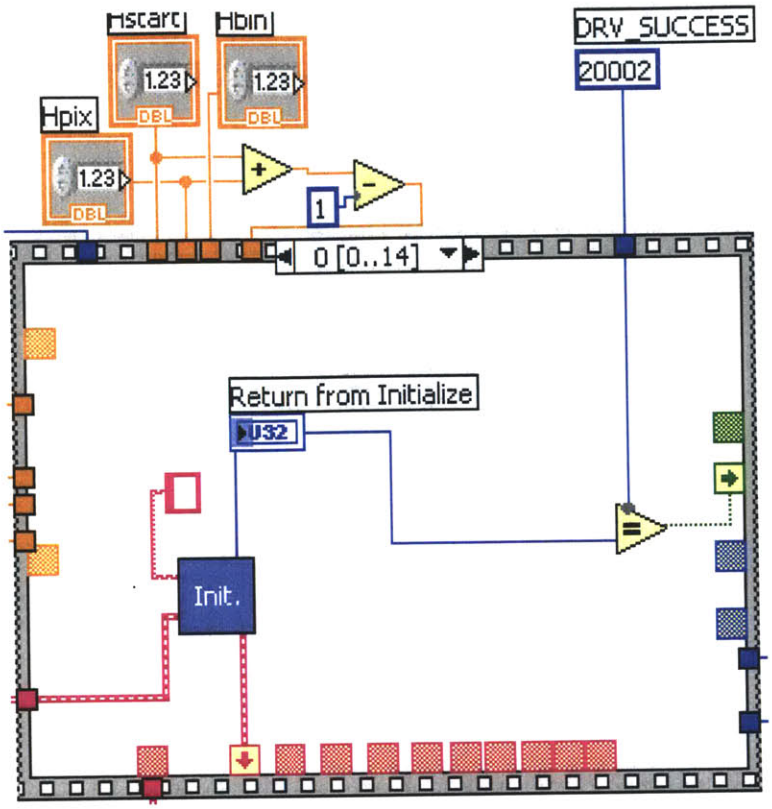

Figure C-3: Frame 0 of CCD camera initialization block: Call initialization function of the Andor Clara CCD camera from the Andor Clara Labview Standard Development Kit. (Requires Labview version 7.0 or higher.)

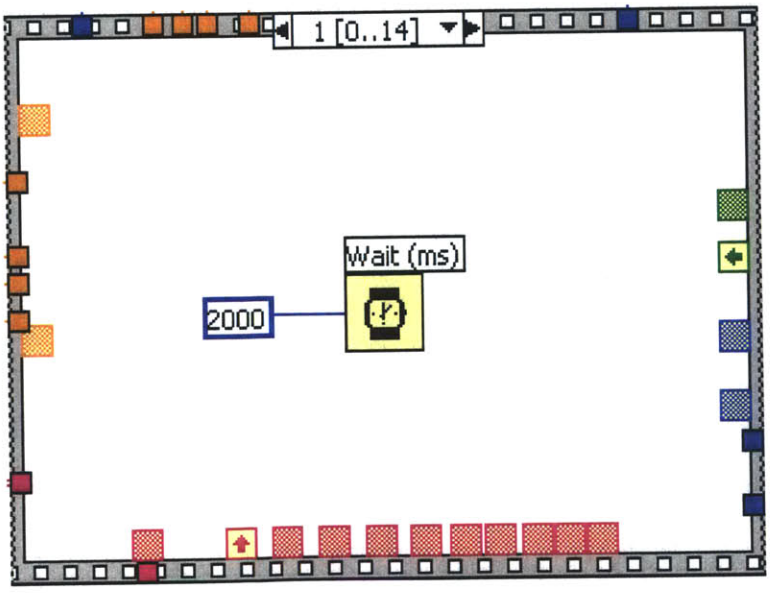

Figure C-4: Frame 1 of CCD camera initialization block: Wait 2 seconds for the initialization function to run.

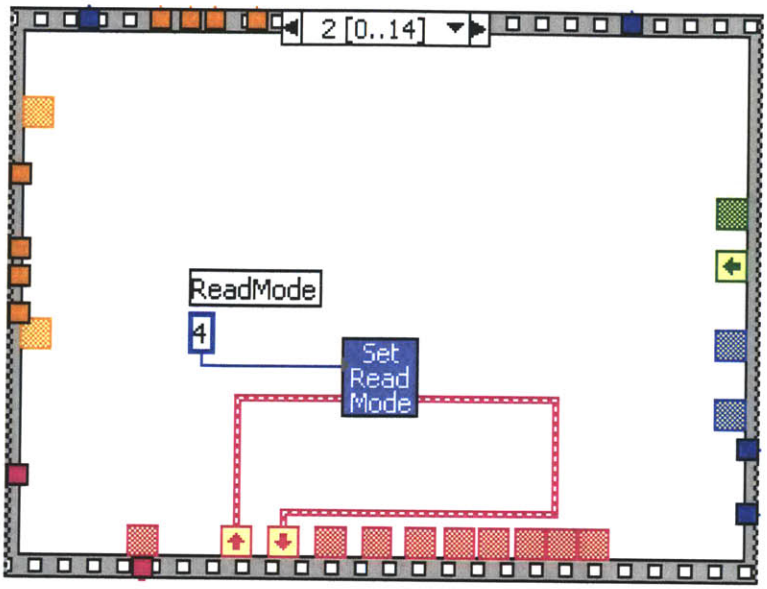

Figure C-5: Frame 2 of CCD camera initialization block: Set Read Mode to 4 = image.

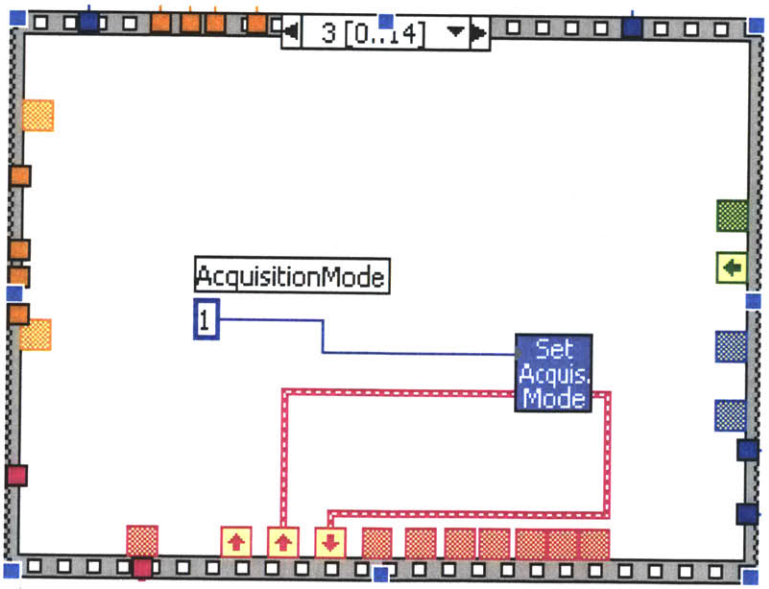

Figure C-6: Frame 3 of CCD camera initialization block: Set Acquisition Mode to 1 = Single Scan.

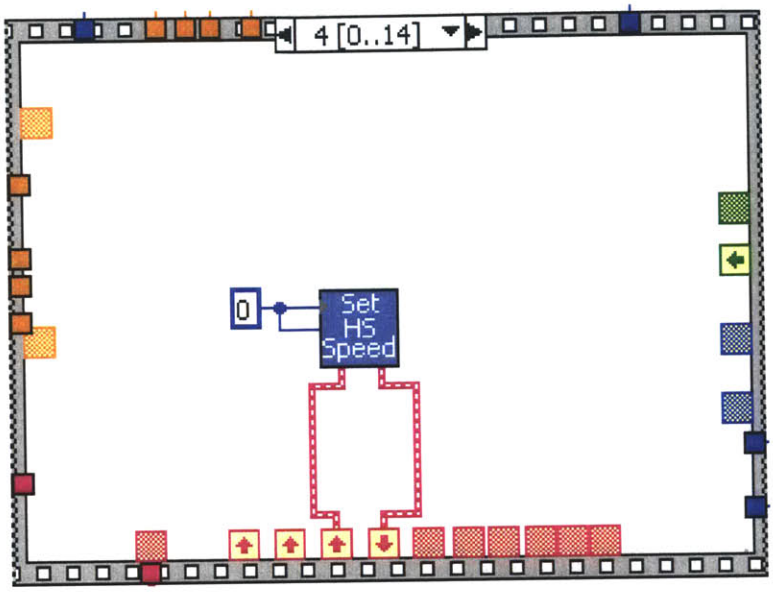

Figure C-7: Frame 4 of CCD camera initialization block: Set the speed at which the pixels are shifted into the output node during readout phase (horizontal speed) of an acquisition to 0 = electron multiplication.

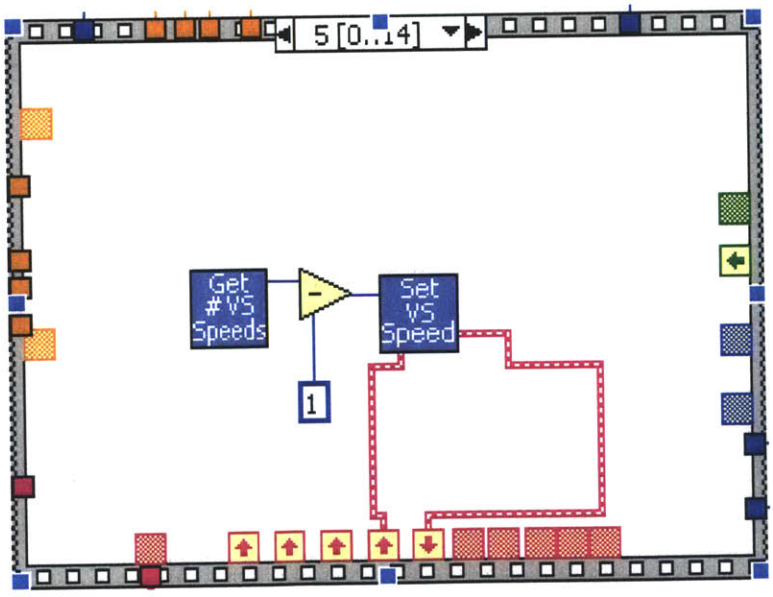

Figure C-8: Frame 5 of CCD camera initialization block: Set vertical speed to its maximum value.

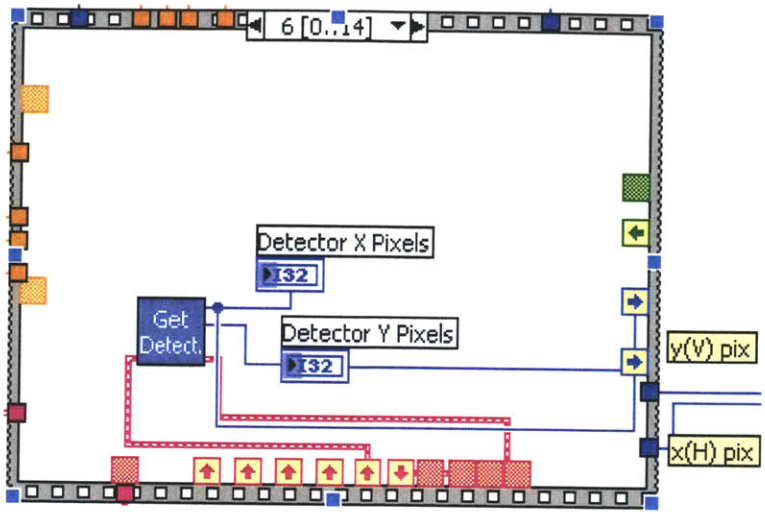

Figure C-9: Frame 6 of CCD camera initialization block: Get the size of the detector in pixels. The horizontal axis is the axis parallel to the readout register and is also the x axis. The vertical axis is the y axis.

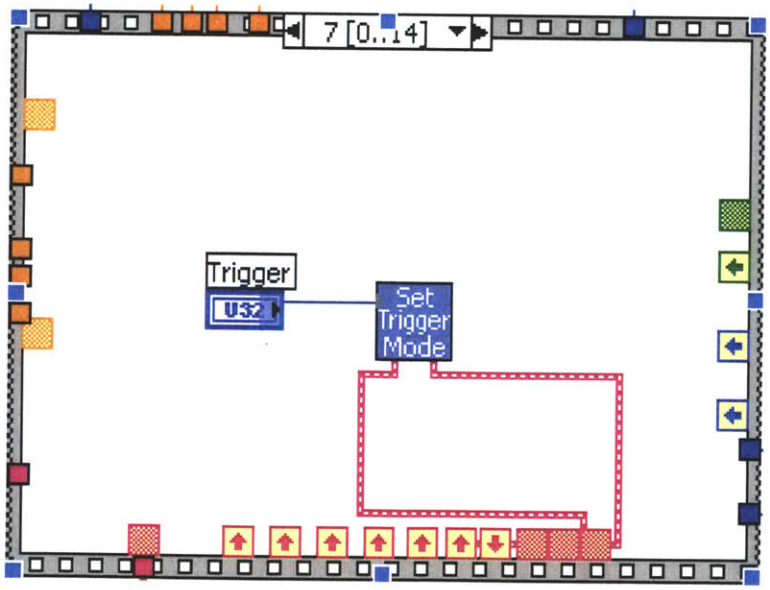

Figure C-10: Frame 7 of CCD camera initialization block: Set trigger mode as a control variable (value = internal for all cases in this thesis).

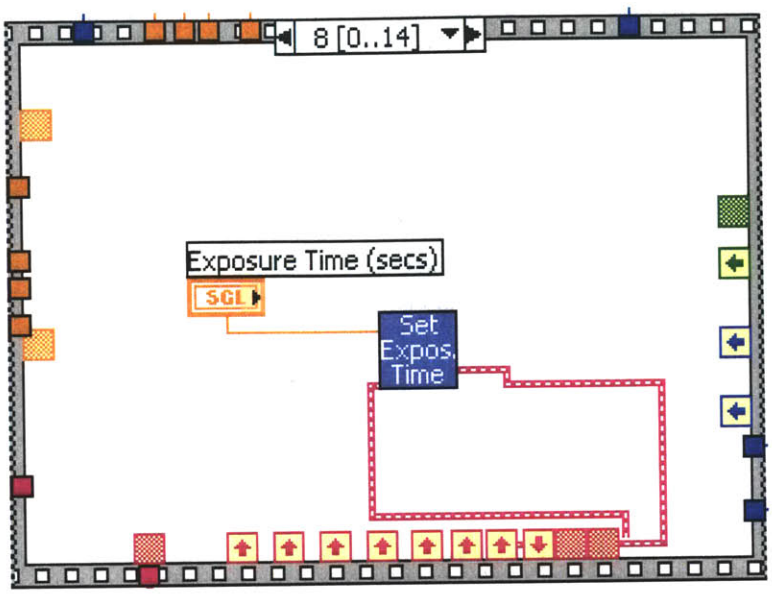

Figure C-11: Frame 8 of CCD camera initialization block: Set Exposure Time as a control variable.

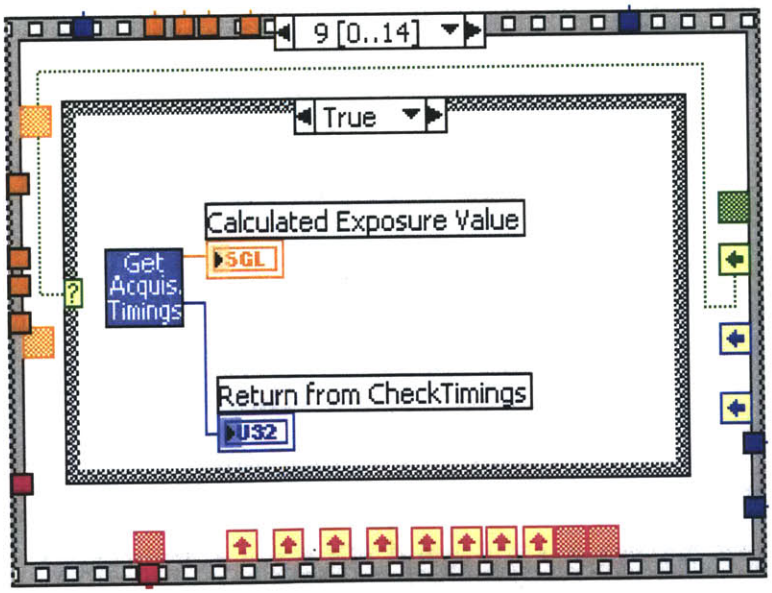

Figure C-12: Frame 9 of CCD camera initialization block: Get the actual exposure times as some input exposure values may not be valid.

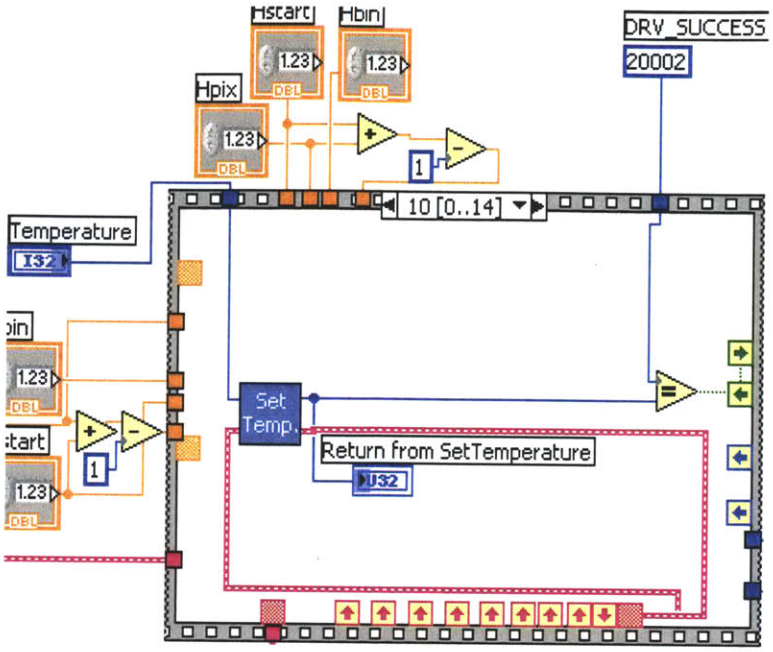

Figure C-13: Frame 10 of CCD camera initialization block: Set temperature of CCD camera to a control variable, Temperature.

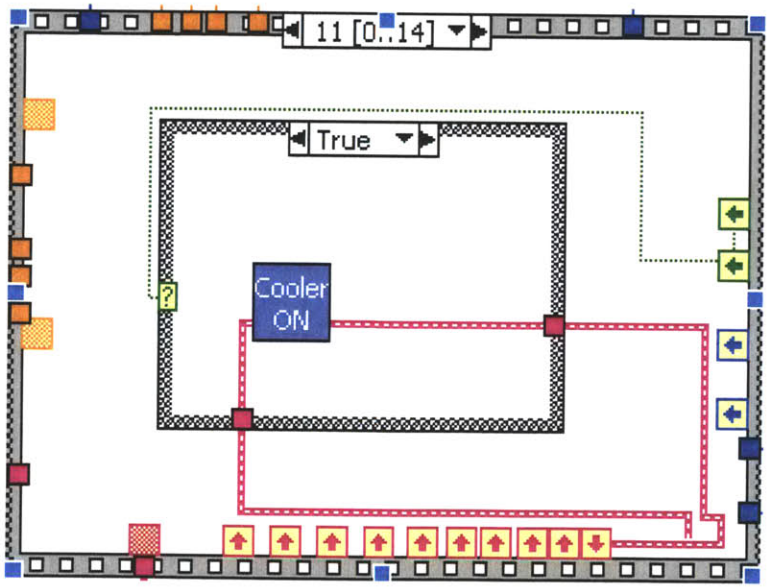

Figure C-14: Frame 11 of CCD camera initialization block: Turn on cooler to reach the set temperature.

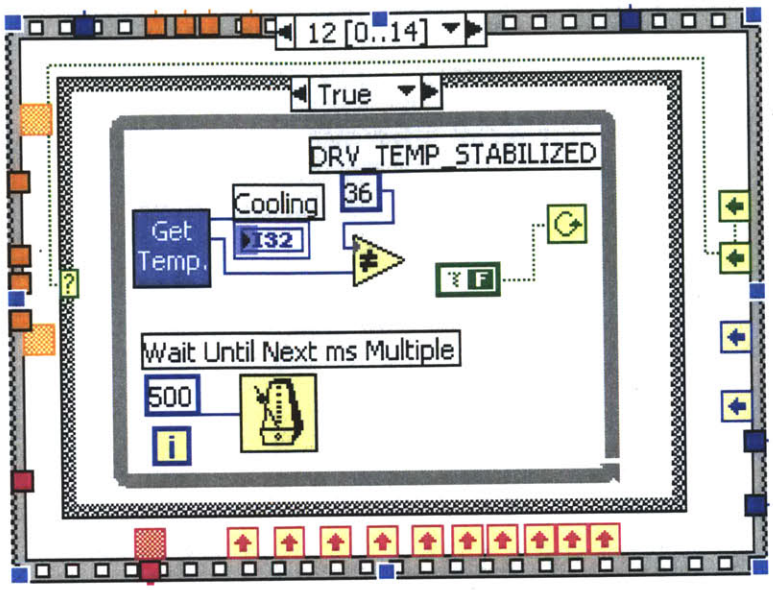

Figure C-15: Frame 12 of CCD camera initialization block: This block can be modified to wait until temperature stabilizes before acquiring images. The block shown below only waits no more than 0.5 seconds and omits this stabilization step for quick initialization. Temperature is stabilized when the returned value for Cooling is 36.

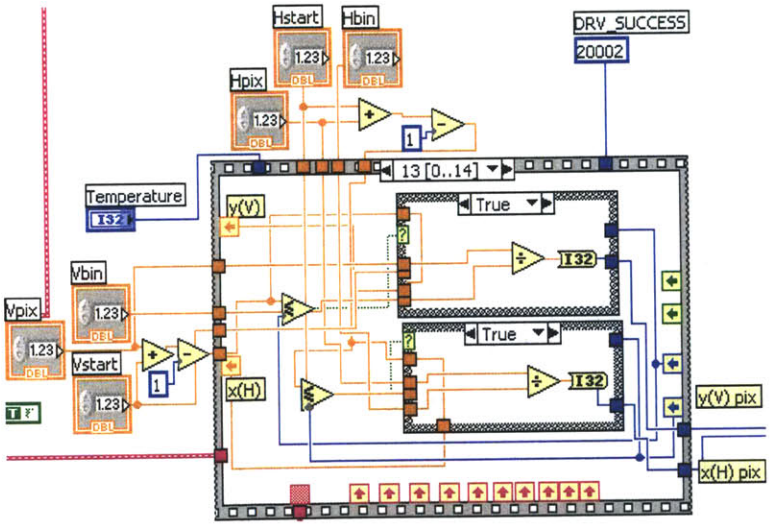

Figure C-16: Frame 13 of CCD camera initialization block: Calculate number of pixels in x and y according to the input x-, y- pixel numbers and binning numbers. This block also checks if the end values of the desired acquisition range exceeds the maximum pixel numbers in x and y.

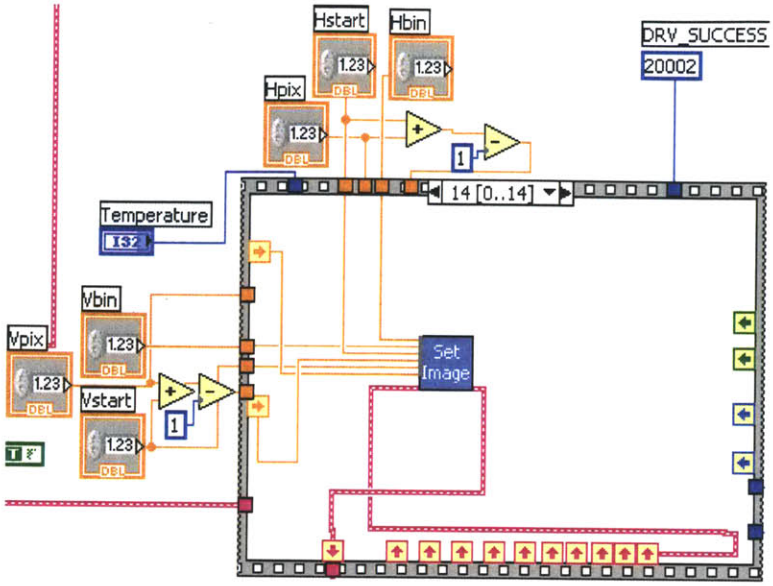

Figure C-17: Frame 14 of CCD camera initialization block: Set imaging rage according to input values.

C.3 Image Acquisition

Figure C-18-C-24 shows the code for stage control and image acquisition.

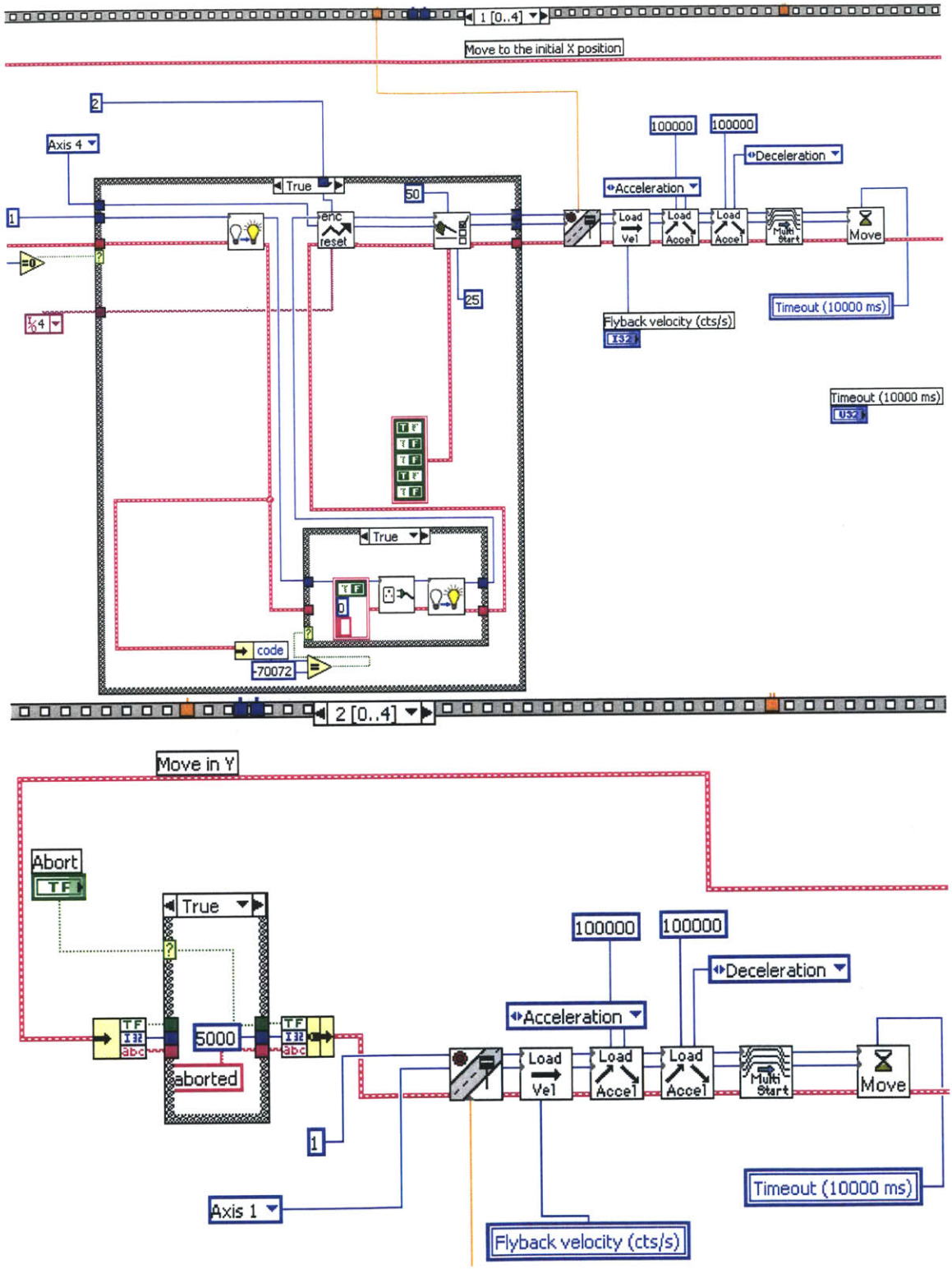

Figure C-18: Initialize PI state control and move to the initial X and Y positions with VIs in the motion library.

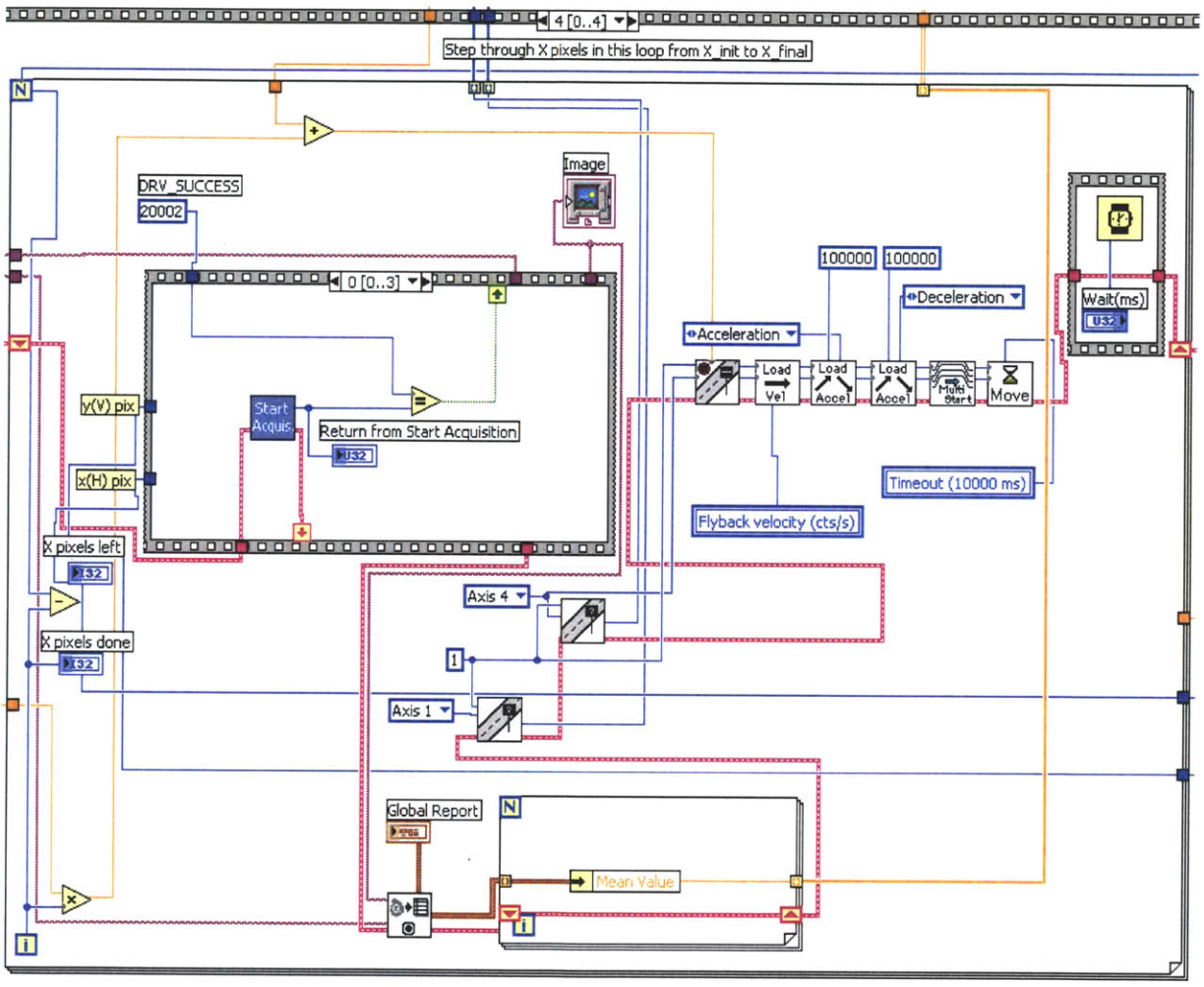

Figure C-19: Acquire image and step X forward after each image.

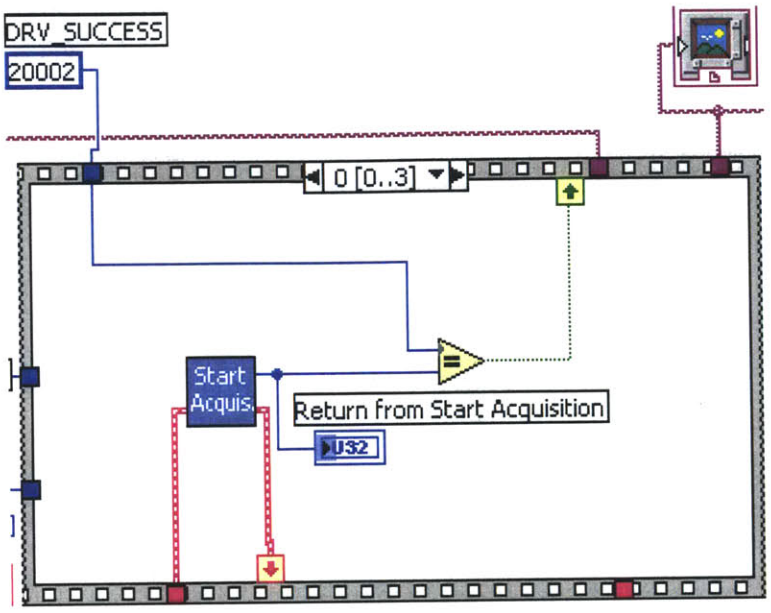

Figure C-20: Frame 0 of image acquisition: Start Acquisition.

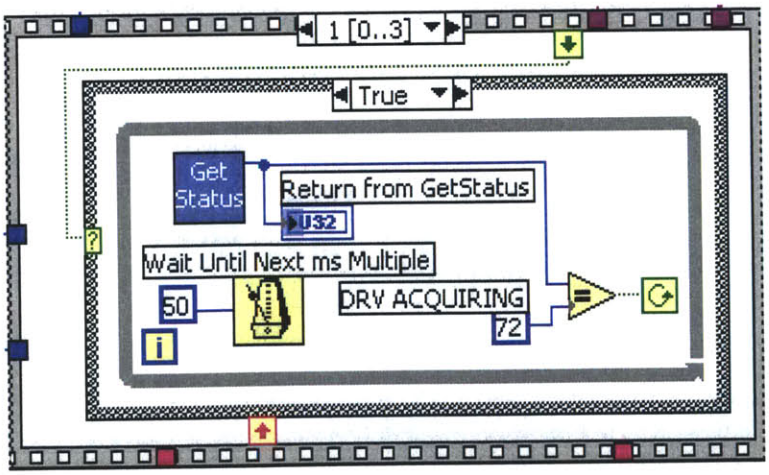

Figure C-21: Frame 1 of image acquisition: Wait until acquisition is complete. Check acquisition status every 50 ms.

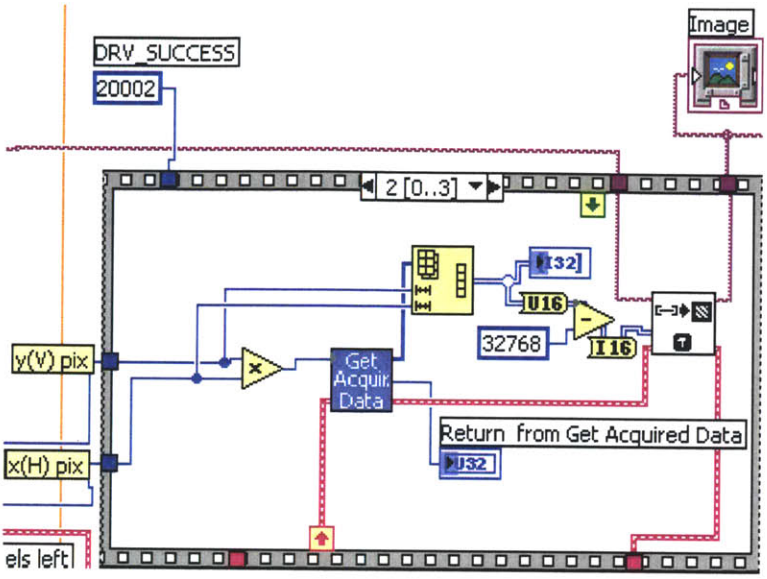

Figure C-22: Frame 2 of image acquisition: Get acquired data with the x-, y- pixel size calculated in the initialization block. Convert the 1-D stream of signed 32-bit integers into a 2-D matrix of unsigned 16-bit integers and then to signed 16-bit integers for ArrayToImage conversion. The image is then sent to another block for extraction of average intensities in the regions of interest.

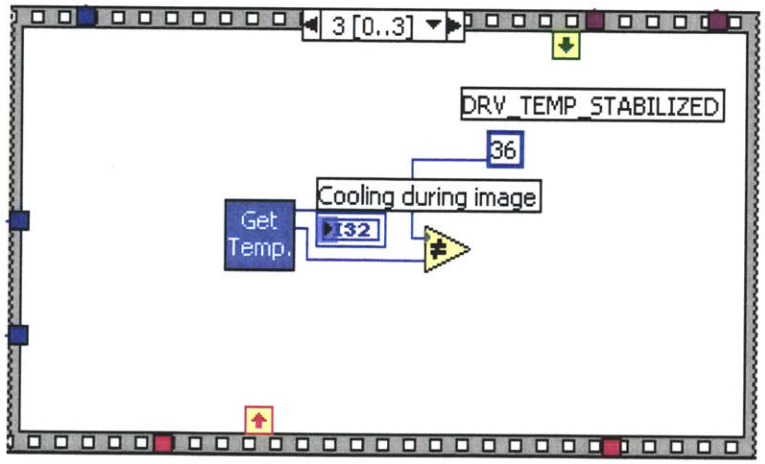

Figure C-23: Frame 3 of image acquisition: Monitor temperature stability of the camera after each acquisition. Camera temperature is stabilized when "Cooling during image" is 36.

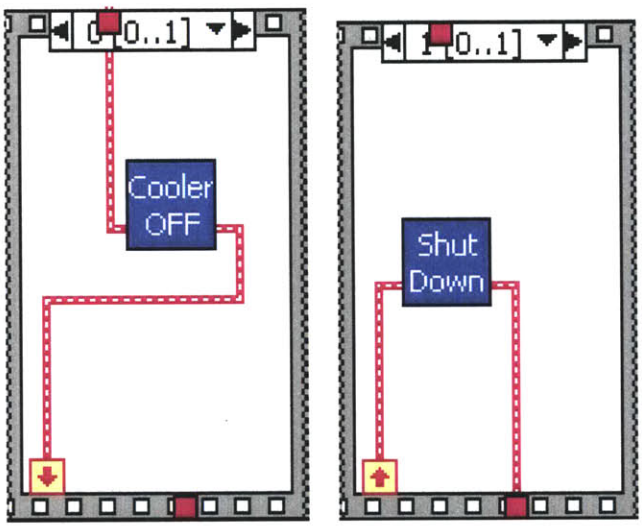

Figure C-24: Turn off camera cooler and shut down CCD camera before exiting the Labview VI.

References

[1] F. Huo, G. Zheng, X. Liao, L.R. Giam, J. Chai, X. Chen, W. Shim, and C.A. Mirkin. Beam pen lithography. *Nature Nanotechnology*, 5:637–640, 2010.

[2] RH French, V. Liberman, HV Tran, J. Feldman, DJ Adelman, RC Wheland, W. Qiu, SJ McLain, O. Nagao, M. Kaku, et al. High-index immersion lithography with second-generation immersion fluids to enable numerical aperatures of 1.55 for cost effective 32-nm half pitches. In *Proceedings of SPIE*, volume 6520, page 65201O, 2007.

[3] J.K.W. Yang and K.K. Berggren. Using high-contrast salty development of hydrogen silsesquioxane for sub-10-nm half-pitch lithography. *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, 25:2025, 2007.

[4] S. Sun and G.J. Leggett. Matching the resolution of electron beam lithography by scanning near-field photolithography. *Nano Letters*, 4(8):1381–1384, 2004.

[5] D. Pires, J.L. Hedrick, A. De Silva, J. Frommer, B. Gotsmann, H. Wolf, M. Despont, U. Duerig, and A.W. Knoll. Nanoscale Three-Dimensional Patterning of Molecular Resists by Scanning Probes. *Science*, 328(5979):732, 2010.

[6] D.S. Ginger, H. Zhang, and C.A. Mirkin. The Evolution of Dip-Pen Nanolithography. *Angewandte Chemie International Edition*, 43(1):30–45, 2004.

[7] K. Salaita, Y. Wang, J. Fragala, R.A. Vega, C. Liu, and C.A. Mirkin. Massively Parallel Dip–Pen Nanolithography with 55 000-Pen Two-Dimensional Arrays. *Angewandte Chemie*, 118(43):7378–7381, 2006.

[8] F. Huo, Z. Zheng, G. Zheng, L.R. Giam, H. Zhang, and C.A. Mirkin. Polymer pen lithography. *Science*, 321(5896):1658, 2008.

[9] Z. Zheng, W.L. Daniel, L.R. Giam, F. Huo, A.J. Senesi, G. Zheng, and C.A. Mirkin. Multiplexed Protein Arrays Enabled by Polymer Pen Lithography: Addressing the Inking Challenge. *Angewandte Chemie*, 121(41):7762–7765, 2009.

[10] L. Huang, A.B. Braunschweig, W. Shim, L. Qin, J.K. Lim, S.J. Hurst, F. Huo, C. Xue, J.W. Jang, and C.A. Mirkin. Matrix-Assisted Dip-Pen Nanolithography and Polymer Pen Lithography. *Small*, 6(10):1077–1081, 2010.

[11] J.Y. Cheng, C.A. Ross, H.I. Smith, and E.L. Thomas. Templated Self-Assembly of Block Copolymers: Top-Down Helps Bottom-Up. *Advanced Materials*, 18(19):2505–2521, 2006.

[12] I. Bita, J.K.W. Yang, Y.S. Jung, C.A. Ross, E.L. Thomas, and K.K. Berggren. Graphoepitaxy of self-assembled block copolymers on two-dimensional periodic patterned templates. *Science*, 321(5891):939, 2008.

[13] J.K.W. Yang, Y.S. Jung, J.B. Chang, RA Mickiewicz, A. Alexander-Katz, C.A. Ross, and K.K. Berggren. Complex self-assembled patterns using sparse commensurate templates with locally varying motifs. *Nature Nanotechnology*, 2010.

[14] S.Y. Chou, P.R. Krauss, and P.J. Renstrom. Imprint of sub-25 nm vias and trenches in polymers. *Applied Physics Letters*, 67:3114, 1995.

[15] J.W. Kingsley, S.K. Ray, A.M. Adawi, G.J. Leggett, and D.G. Lidzey. Optical nanolithography using a scanning near-field probe with an integrated light source. *Applied Physics Letters*, 93:213103, 2008.

[16] G.J. Leggett. Scanning near-field photolithography-surface photochemistry with nanoscale spatial resolution. *Chemical Society Reviews*, 35(11):1150–1161, 2006.

[17] S. Kwon, W. Chang, and S. Jeong. Shape and size variations during nanopatterning of photoresist using near-field scanning optical microscope. *Ultramicroscopy*, 105(1-4):316–323, 2005.

[18] J.G. Goodberlet and H. Kavak. Patterning sub-50 nm features with near-field embedded-amplitude masks. *Applied Physics Letters*, 81:1315, 2002.

[19] T. Ito, T. Yamada, Y. Inao, T. Yamaguchi, N. Mizutani, and R. Kuroda. Fabrication of half-pitch 32 nm resist patterns using near-field lithography with a-Si mask. *Applied Physics Letters*, 89:033113, 2006.

[20] H. Raether. *Surface plasmons*. Springer-Verlag New York, 1988.

[21] W.L. Barnes, A. Dereux, and T.W. Ebbesen. Surface plasmon subwavelength optics. *Nature*, 424(6950):824–830, 2003.

[22] E.X. Jin and X. Xu. Obtaining super resolution light spot using surface plasmon assisted sharp ridge nanoaperture. *Applied Physics Letters*, 86:111106, 2005.

[23] L. Wang, S.M. Uppuluri, E.X. Jin, and X. Xu. Nanolithography using high transmission nanoscale bowtie apertures. *Nano lett*, 6(3):361–364, 2006.

[24] L. Wang, E.X. Jin, S.M. Uppuluri, and X. Xu. Contact optical nanolithography using nanoscale C-shaped apertures. *Optics Express*, 14(21):9902–9908, 2006.

[25] Y. Wang, W. Srituravanich, C. Sun, and X. Zhang. Plasmonic nearfield scanning probe with high transmission. *Nano letters*, 8(9):3041–3045, 2008.

[26] W. Srituravanich, L. Pan, Y. Wang, C. Sun, D.B. Bogy, and X. Zhang. Flying plasmonic lens in the near field for high-speed nanolithography. *Nature Nanotechnology*, 3(12):733–737, 2008.

[27] Y. Kim, S. Kim, H. Jung, E. Lee, and J.W. Hahn. Plasmonic nano lithography with a high scan speed contact probe. *Optics Express*, 17(22):19476–19485, 2009.

[28] C.N. LaFratta, J.T. Fourkas, T. Baldacchini, and R.A. Farrer. Multiphoton fabrication. *Angewandte Chemie International Edition*, 46(33):6238–6258, 2007.

[29] G. Witzgall, R. Vrijen, E. Yablonovitch, V. Doan, and B.J. Schwartz. Single-shot two-photon exposure of commercial photoresist for the production of three-dimensional structures. *Optics letters*, 23(22):1745–1747, 1998.

[30] D. Tan, Y. Li, F. Qi, H. Yang, Q. Gong, X. Dong, and X. Duan. Reduction in feature size of two-photon polymerization using SCR500. *Applied Physics Letters*, 90:071106, 2007.

[31] W. Haske, V.W. Chen, J.M. Hales, W. Dong, S. Barlow, S.R. Marder, and J.W. Perry. 65 nm feature sizes using visible wavelength 3-D multiphoton lithography. *Optics Express*, 15(6):3426–3436, 2007.

[32] S. Kawata, H.B. Sun, T. Tanaka, and K. Takada. Finer features for functional microdevices. *Nature*, 412(6848):697–698, 2001.

[33] J.F. Xing, X.Z. Dong, W.Q. Chen, X.M. Duan, N. Takeyasu, T. Tanaka, and S. Kawata. Improving spatial resolution of two-photon microfabrication by using photoinitiator with high initiating efficiency. *Applied Physics Letters*, 90:131106, 2007.

[34] X. Yin, N. Fang, X. Zhang, I.B. Martini, and B.J. Schwartz. Near-field two-photon nanolithography using an apertureless optical probe. *Applied Physics Letters*, 81:3663, 2002.

[35] N. Murphy-DuBay, L. Wang, EC Kinzel, SM Uppuluri, and X. Xu. Nanopatterning using NSOM probes integrated with high transmission nanoscale bowtie aperture. *Optics express*, 16(4):2584, 2008.

[36] AP Chu, KK Berggren, KS Johnson, and MG Prentiss. A virtual slit for atom optics and nanolithography. *Quantum Semiclass. Opt*, 8:521, 1996.

[37] KS Johnson et al. Localization of Metastable Atom Beams with Optical Standing Waves: Nanolithography at the Heisenberg Limit. *Science*, 280(5369):1583–1586, 1998.

[38] S.W. Hell and J. Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett*, 19(11):780–782, 1994.

[39] KK Berggren, A. Bard, JL Wilbur, JD Gillaspy, AG Helg, JJ McClelland, SL Rolston, WD Phillips, M. Prentiss, and GM Whitesides. Microlithography by using neutral metastable atoms and self-assembled monolayers. *Science*, 269(5228):1255, 1995.

[40] R. Menon and H.I. Smith. Absorbance-modulation optical lithography. *JOSA A*, 23(9):2290–2294, 2006.

[41] L. Li, R.R. Gattass, E. Gershgoren, H. Hwang, and J.T. Fourkas. Achieving {lambda}/20 Resolution by One-Color Initiation and Deactivation of Polymerization. *Science*, 324(5929):910, 2009.

[42] J.O. Hirschfelder, C.F. Curtiss, R.B. Bird, and University of Wisconsin. Theoretical Chemistry Laboratory. *Molecular theory of gases and liquids*, volume 26. Wiley New York, 1954.

[43] M.P. Stocker, L. Li, R.R. Gattass, and J.T. Fourkas. Multiphoton photoresists giving nanoscale resolution that is inversely dependent on exposure time. *Nature Chemistry*, 2011.

[44] T.F. Scott, B.A. Kowalski, A.C. Sullivan, C.N. Bowman, and R.R. McLeod. Two-color single-photon photoinitiation and photoinhibition for subdiffraction photolithography. *Science*, 324(5929):913, 2009.

[45] M. Born, E. Wolf, and AB Bhatia. *Principles of optics*, volume 508. Pergamon press Oxford, 1975.

[46] P. Bourgeat, F. Meriaudeau, K.W. Tobin, and P. Gorria. Content based segmentation of patterned wafers. *Journal of Electronic Imaging*, 13:428, 2004.

[47] S. Teuber, A. Bzdurek, AC Durr, J. Heumann, and C. Holfeld. Limitations of optical reticle inspection for 45-nm node and beyond [6349-29]. In *PROCEEDINGS-SPIE THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING*, volume 6349, page 63490. International Society for Optical Engineering; 1999, 2006.

[48] R. Peters. Checking and fixing the cellular nanomachinery: towards medical nanoscopy. *Trends in molecular medicine*, 12(2):83–89, 2006.

[49] S.W. Hell. Far-field optical nanoscopy. *Science*, 316(5828):1153, 2007.

[50] SW Hell and M. Kroug. Ground-state-depletion fluorscence microscopy: A concept for breaking the diffraction resolution limit. *Applied Physics B: Lasers and Optics*, 60(5):495–497, 1995.

[51] S. Bretschneider, C. Eggeling, and S.W. Hell. Breaking the diffraction barrier in fluorescence microscopy by optical shelving. *Physical review letters*, 98(21):218103, 2007.

[52] V. Westphal and S.W. Hell. Nanoscale resolution in the focal plane of an optical microscope. *Physical review letters*, 94(14):143903, 2005.

[53] M. Hofmann, C. Eggeling, S. Jakobs, and S.W. Hell. Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 102(49):17565, 2005.

[54] H. Shroff, C.G. Galbraith, J.A. Galbraith, and E. Betzig. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. *Nature Methods*, 5(5):417–423, 2008.

[55] E. Betzig, G.H. Patterson, R. Sougrat, O.W. Lindwasser, S. Olenych, J.S. Bonifacino, M.W. Davidson, J. Lippincott-Schwartz, and H.F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642, 2006.

[56] S.T. Hess, T.P.K. Girirajan, and M.D. Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical journal*, 91(11):4258–4272, 2006.

[57] A. Egner, C. Geisler, C. Von Middendorff, H. Bock, D. Wenzel, R. Medda, M. Andresen, A.C. Stiel, S. Jakobs, C. Eggeling, et al. Fluorescence nanoscopy in whole cells by asynchronous localization of photoswitching emitters. *Biophysical journal*, 93(9):3285–3290, 2007.

[58] B. Huang, W. Wang, M. Bates, and X. Zhuang. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*, 319(5864):810, 2008.

[59] M.J. Rust, M. Bates, and X. Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature methods*, 3(10):793–796, 2006.

[60] M. Bates, B. Huang, G.T. Dempsey, and X. Zhuang. Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science*, 317(5845):1749, 2007.

[61] A. Yildiz, J.N. Forkey, S.A. McKinney, T. Ha, Y.E. Goldman, and P.R. Selvin. Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science*, 300(5628):2061, 2003.

[62] M.F. Juette, T.J. Gould, M.D. Lessard, M.J. Mlodzianoski, B.S. Nagpure, B.T. Bennett, S.T. Hess, and J. Bewersdorf. Three-dimensional sub–100 nm resolution fluorescence microscopy of thick samples. *Nature Methods*, 5(6):527–529, 2008.

[63] S.R.P. Pavani, M.A. Thompson, J.S. Biteen, S.J. Lord, N. Liu, R.J. Twieg, R. Piestun, and WE Moerner. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proceedings of the National Academy of Sciences*, 106(9):2995, 2009.

[64] G. Shtengel, J.A. Galbraith, C.G. Galbraith, J. Lippincott-Schwartz, J.M. Gillette, S. Manley, R. Sougrat, C.M. Waterman, P. Kanchanawong, M.W. Davidson, et al. Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure. *Proceedings of the National Academy of Sciences*, 106(9):3125, 2009.

[65] J. Bierwagen, I. Testa, J. Fölling, D. Wenzel, S. Jakobs, C. Eggeling, and S.W. Hell. Far-Field Autofluorescence Nanoscopy. *Nano Letters*, pages 175–197.

[66] J.W. Goodman. *Introduction to Fourier optics*. Roberts & Company Publishers, 2005.

[67] R. Heintzmann and C.G. Cremer. Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating. In *Proceedings of SPIE*, volume 3568, page 185, 1999.

[68] M.G.L. Gustafsson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy*, 198(2):82–87, 2000.

[69] L. Shao, B. Isaac, S. Uzawa, D.A. Agard, J.W. Sedat, and M.G.L. Gustafsson. I5S: wide-field light microscopy with 100-nm-scale resolution in three dimensions. *Biophysical journal*, 94(12):4971–4983, 2008.

[70] R. Heintzmann and M.G.L. Gustafsson. Subdiffraction resolution in continuous samples. *Nature Photonics*, 3(7):362–364, 2009.

[71] M.G.L. Gustafsson. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 102(37):13081, 2005.

[72] K. K
"onig. Multiphoton microscopy in life sciences. *Journal of Microscopy*, 200(2):83–104, 2000.

[73] F. Helmchen and W. Denk. Deep tissue two-photon microscopy. *Nature methods*, 2(12):932–940, 2005.

[74] M. Albota, D. Beljonne, J.L. Brédas, J.E. Ehrlich, J.Y. Fu, A.A. Heikal, S.E. Hess, T. Kogej, M.D. Levin, S.R. Marder, et al. Design of organic molecules with large two-photon absorption cross sections. *Science*, 281(5383):1653, 1998.

[75] W. Zhou, S.M. Kuebler, K.L. Braun, T. Yu, J.K. Cammack, C.K. Ober, J.W. Perry, and S.R. Marder. An efficient two-photon-generated photoacid applied to positive-tone 3D microfabrication. *Science*, 296(5570):1106, 2002.

[76] E. Betzig and R.J. Chichester. Single molecules observed by near-field scanning optical microscopy. *Science*, 262(5138):1422, 1993.

[77] E.J. Sánchez, L. Novotny, and X.S. Xie. Near-field fluorescence microscopy based on two-photon excitation with metal tips. *Physical Review Letters*, 82(20):4014–4017, 1999.

[78] H. Yokota, K. Saito, and T. Yanagida. Single molecule imaging of fluorescently labeled proteins on metal by surface plasmons in aqueous solution. *Physical Review Letters*, 80(20):4606–4609, 1998.

[79] F. Yu, B. Persson, S. L
"ofås, and W. Knoll. Attomolar sensitivity in bioassays based on surface plasmon fluorescence spectroscopy. *J. Am. Chem. Soc*, 126(29):8902–8903, 2004.

[80] K. Tawa and K. Morigaki. Substrate-supported phospholipid membranes studied by surface plasmon resonance and surface plasmon fluorescence spectroscopy. *Biophysical journal*, 89(4):2750–2758, 2005.

[81] K. Kim, Y. Oh, W. Lee, and D. Kim. Plasmonics-based spatially activated light microscopy for super-resolution imaging of molecular fluorescence.

[82] S. Kawata, Y. Inouye, and P. Verma. Plasmonics for near-field nano-imaging and superlensing. *Nature Photonics*, 3(7):388–394, 2009.

[83] JB Pendry. Negative Refraction Makes a Perfect Lens. *Physical Review Letters*, 85(18):3966–3969, 2000.

[84] T. Taubner, D. Korobkin, Y. Urzhumov, G. Shvets, and R. Hillenbrand. Near-field microscopy through a SiC superlens. *Science*, 313(5793):1595, 2006.

[85] N. Fang and X. Zhanga. Imaging properties of a metamaterial superlens. *APPLIED PHYSICS LETTERS*, 82(2), 2003.

[86] N. Fang, H. Lee, C. Sun, and X. Zhang. Sub-diffraction-limited optical imaging with a silver superlens. *Science*, 308(5721):534, 2005.

[87] Z. Liu, H. Lee, Y. Xiong, C. Sun, and X. Zhang. Far-field optical hyperlens magnifying sub-diffraction-limited objects. *Science*, 315(5819):1686, 2007.

[88] G. Shvets, S. Trendafilov, JB Pendry, and A. Sarychev. Guiding, focusing, and sensing on the subwavelength scale using metallic wire arrays. *Physical review letters*, 99(5):53903, 2007.

[89] II Smolyaninov, YJ Hung, and CC Davis. Magnifying superlens in the visible frequency range. *Science (New York, NY)*, 315(5819):1699, 2007.

[90] D. Gil, R. Menon, DJD Carter, and H.I. Smith. Lithographic patterning and confocal imaging with zone plates. *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, 18:2881, 2000.

[91] B. Harke, J. Keller, CK Ullal, V. Westphal, A. Sch "onle, and SW Hell. Resolution scaling in STED microscopy. *Optics express*, 16(6):4154, 2008.

[92] Rajesh Menon and Henry I. Smith. Absorbance-modulation optical lithography. *J. Opt. Soc. Am. A*, 23(9):2290, 2006.

[93] SR Arridge, M. Schweiger, M. Hiraoka, and DT Delpy. A finite element approach for modeling photon transport in tissue. *MEDICAL PHYSICS-LANCASTER PA-*, 20:299–299, 1993.

[94] C. Trundle. Contrast enhanced photolithography, July 17 1990. US Patent 4,942,113.

[95] A. Taflove, S.C. Hagness, et al. *Computational electrodynamics: The finite-difference time-domain method.* Artech House Norwood, MA, 2000.

[96] Rajesh Menon, Hsin-Yu Tsai, and Samuel W. Thomas III. Far-Field Generation of Localized Light Fields using Absorbance Modulation. *Physical Review Letters*, 98:043905, 2007.

[97] H.Y. Tsai, G.M. Wallraff, and R. Menon. Spatial-frequency multiplication via absorbance modulation. *Applied Physics Letters*, 91:094103, 2007.

[98] P.J. Darcy, H.G. Heller, P.J. Strydom, and J. Whittall. Photochromic heterocyclic fulgides. Part 2. Electrocyclic reactions of (E)-α-2, 5-dimethyl-3-furylethylidene (alkyl-substituted methylene) succinic anhydrides. *Journal of the Chemical Society, Perkin Transactions 1*, 1981:202–205, 1981.

[99] Y. Yokoyama. Fulgides for memories and switches. *Chem. Rev*, 100(5):1717–1740, 2000.

[100] M. Irie. Diarylethenes for memories and switches. *Chem. Rev*, 100(5):1685–1716, 2000.

[101] S. Leea, J. rey Byersb, K. Jena, P. Zimmermanb, B. Riceb, N.J. Turroc, and C.G. Willsona. An analysis of double exposure lithography options. In *Proc. of SPIE Vol*, volume 6924, pages 69242A–1.

[102] R. Bristol, D. Shykind, S. Kim, Y. Borodovsky, E. Schwartz, C. Turner, G. Masson, K. Min, K. Esswein, J.M. Blackwell, et al. Double-exposure materials for pitch division with 193nm lithography: requirements, results. In *Proceedings of SPIE*, volume 7273, page 727307, 2009.

[103] V. Westphal and S.W. Hell. Nanoscale Resolution in the Focal Plane of an Optical Microscope. *Physical Review Letters*, 94(14):143903, 2005.

[104] K.I. Willig, R.R. Kellner, R. Medda, B. Hein, S. Jakobs, and S.W. Hell. Nanoscale resolution in GFP-based microscopy. *Nature Methods*, 3(9):721–723, 2006.

[105] MP MacDonald, L. Paterson, G. Armstrong, J. Arlt, P. Bryant, W. Sibbett, and K. Dholakia. Laguerre-Gaussian laser modes for biophotonics and micromanipulation. In *Proceedings of SPIE*, volume 5147, page 48, 2003.

[106] A. Vasara, J. Turunen, and A.T. Friberg. Realization of general nondiffracting beams with computer-generated holograms. *Journal of the Optical Society of America A*, 6(11):1748–1754, 1989.

[107] C. Paterson and R. Smith. Higher-order Bessel waves produced by axicon-type computer-generated holograms. *Optics Communications*, 124(1-2):121–130, 1996.

[108] H.I. Smith, R. Menon, A. Patel, D. Chao, M. Walsh, and G. Barbastathis. Zone-plate-array lithography: A low-cost complement or competitor to scanning-electron-beam lithography. *Microelectronic Engineering*, 83(4-9):956–961, 2006.

[109] T. Watanabe, M. Fujii, Y. Watanabe, N. Toyama, and Y. Iketaki. Generation of a doughnut-shaped beam using a spiral phase plate. *Review of Scientific Instruments*, 75:5131, 2004.

[110] H.Y. Tsai, H.I. Smith, and R. Menon. Fabrication of spiral-phase diffractive elements using scanning-electron-beam lithography. *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, 25:2068, 2007.

[111] Y. Arieli, S. Noach, S. Ozeri, and N. Eisenberg. Design of diffractive optical elements for multiple wavelengths. *Applied optics*, 37(26):6174–6177, 1998.

[112] D.W. Sweeney and G.E. Sommargren. Harmonic diffractive lenses. *Applied Optics*, 34(14):2469–2475, 1995.

[113] D. Faklis and G.M. Morris. Spectral properties of multiorder diffractive lenses. *Applied Optics*, 34(14):2462–2468, 1995.

[114] G.T. Di Francia. Super-gain antennas and optical resolving power. *Il Nuovo Cimento (1943-1954)*, 9:426–438, 1952.

[115] M. Mitchell. *An introduction to genetic algorithms*. The MIT press, 1998.

[116] Inc. The MathWorks. Documentation for genetic algorithm and direct search toolbox in matlab. `http:www.mathworks.com/access/helpdesk/help/toolbox/gads/`.

[117] R. Menon, P. Rogge, and H.Y. Tsai. Design of diffractive lenses that generate optical nulls without phase singularities. *Journal of the Optical Society of America A*, 26(2):297–304, 2009.

[118] Y. Granik. Fast pixel-based mask optimization for inverse lithography. *Journal of Microlithography, Microfabrication, and Microsystems*, 5:043002, 2006.

[119] A. Erdmann, R. Farkas, T. Fuehner, B. Tollkuehn, and G. Kokai. Mask and source optimization for lithographic imaging systems. In *Proceedings of SPIE*, 2003.

[120] T. Fuhner and A. Erdmann. Improved mask and source representations for automatic optimization of lithographic process conditions using a genetic algorithm. In *Proceedings of SPIE*, volume 5754, page 415, 2005.

[121] D.W. Prather and S. Shi. Formulation and application of the finite-difference time-domain method for the analysis of axially symmetric diffractive optical elements. *JOSA A*, 16(5):1131–1142, 1999.

[122] D. Gil, R. Menon, and H.I. Smith. Fabrication of high-numerical-aperture phase zone plates with a single lithography exposure and no etching. *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, 21:2956, 2003.

[123] R. Menon, D. Gil, and H.I. Smith. Experimental characterization of focusing by high-numerical-aperture zone plates. *Journal of the Optical Society of America A*, 23(3):567–571, 2006.

[124] TA Fulton and GJ Dolan. New approach to electron beam lithography. *Applied Physics Letters*, 42:752, 1983.

[125] GJ Dolan and T. Fulton. Canyon lithography. *IEEE Electron Device Letters*, 4(6):178–180, 1983.

[126] D. Gil, J.T. Hastings, J.G. Goodberlet, R. Menon, D.J. Carter, and H.I. Smith. System and method for fabrication and replication of diffractive optical elements for maskless lithography, March 25 2008. US Patent 7,348,104.

[127] D. Gil. Maskless nanolithography and imaging with diffractive optical arrays. 2003.

[128] T.R.M. Sales and G.M. Morris. Diffractive superresolution elements. *Journal of the Optical Society of America A*, 14(7):1637–1646, 1997.

[129] R. Dorn, S. Quabis, and G. Leuchs. Sharper focus for a radially polarized light beam. *Physical review letters*, 91(23):233901, 2003.

[130] D. Wildanger, J. B
"uckers, V. Westphal, S.W. Hell, and L. Kastrup. A STED microscope aligned by design. *Optics Express*, 17(18):16100–16110, 2009.

[131] V. Westphal, S.O. Rizzoli, M.A. Lauterbach, D. Kamin, R. Jahn, and S.W. Hell. Video-rate far-field optical nanoscopy dissects synaptic vesicle movement. *Science*, 320(5873):246, 2008.